OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: April 5, 2003, 05:23:15 ; Search time 9963 Seconds
(without alignments)
10147.862 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtggactgctatcacccttgg........tgatcgctcgtcgttcctaa 3474

Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 1455140287B residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3474	100.0	3474	6	AX453601	AX453601 Sequence
2	3474	100.0	3474	6	AX453603	AX453603 Sequence
3	3398.8	97.8	3621	6	AR123974	AR123974 Sequence
4	3398.8	97.8	3637	1	AF038548	AF038548 Corynebac
5	3398.8	97.8	3728	1	CGPYC	Y09548 Corynebacte
6	3398.8	97.8	3728	6	A97276	A97276 Sequence 1
7	3398.8	97.8	332050	1	AP005276	AP005276 Corynebac
8	3398.8	97.8	349980	6	AX127145	AX127145 Sequence
9	3394	97.7	3657	6	AX503915	AX503915 Corynebac
10	3344.8	96.3	3420	6	AX120849	AX120849 Sequence
11	2256.2	64.9	4851	6	AB083299	AB083299 Corynebac
12	1564.8	45.0	1719	6	AX064993	AX064993 Sequence
13	1351.8	38.9	1406	6	AX064995	AX064995 Sequence
c 14	1351	38.9	40104	1	SCF11	AL132662 Streptomy
15	1329.8	38.3	4218	1	AF262949	AF262949 Mycobacte
c 16	1283.2	36.9	17992	1	AE007125	AE007125 Mycobacte
c 17	1283.2	36.9	43523	1	MTCY349	Z83018 Mycobacteri
18	1238.4	35.6	31175	1	MTU00024	U00024 Mycobacteri
19	1238.4	35.6	31176	1	MSGY2	AD000009 Mycobacte
20	1020.2	29.4	1083	6	AX064989	AX064989 Sequence
21	1020.2	29.4	1083	6	AX064991	AX064991 Sequence
22	883	25.4	939	6	AX064985	AX064985 Sequence
23	883	25.4	939	6	AX064987	AX064987 Sequence
24	695.6	20.0	3916	8	AF097728	AF097728 Aspergill
25	691.6	19.9	3559	5	AF509529	AF509529 Gallus ga
26	688.6	19.8	3953	9	HSU04641	U04641 Human pyruv
27	688.6	19.8	4058	9	BC011617	BC011617 Homo sapi
28	687	19.8	3960	9	HSU30891	U30891 Human pyruv
29	684.2	19.7	4067	10	MUSMPYR	L09192 Mus musculu
30	682.6	19.6	4024	10	RNU36585	U36585 Rattus norv
31	681	19.6	3945	6	AX401890	AX401890 Sequence
32	681	19.6	3945	10	RNU32314	U32314 Rattus norv
33	674.2	19.4	4017	6	AX410383	AX410383 Sequence
34	674.2	19.4	4017	9	S72370	S72370 pyruvate ca
35	670.4	19.3	6513	8	ANI9972	AJ009972 Aspergill
36	665.2	19.1	4371	3	AF132152	AF132152 Drosophil
37	648	18.7	42926	1	MLCB1243	AL023635 Mycobacte
c 38	648	18.7	348950	1	MLEPRTN7	AL583923 Mycobacte
39	647.8	18.6	38477	1	MSGB1723CS	MSGB1723CS Mycobacteri
40	640.4	18.4	3800	1	REU51439	U51439 Rhizobium e
c 41	640.4	18.4	349116	1	AP003003	AP003003 Mesorhizo
42	639	18.4	4000	5	AF295372	AF295372 Danio rer
43	622	17.9	11295	1	AE009469	AE009469 Brucella
44	611	17.6	3727	1	D83706	D83706 Bacillus st
c 45	609	17.5	7076	1	AE009219	AE009219 Agrobacte

ALIGNMENTS

RESULT 1
AX453601
LOCUS   AX453601   3474 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 1 from Patent WO0231158.
ACCESSION   AX453601
VERSION   AX453601.1 GI:21712841
KEYWORDS   .
SOURCE   Corynebacterium glutamicum.
ORGANISM   Corynebacterium glutamicum
             Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
             Actinomycetales; Corynebacterineae; Corynebacteriaceae;
             Corynebacterium.
REFERENCE   1
AUTHORS   Hanke,P.D.
TITLE   Feedback-resistant pyruvate carboxylase gene from corynebacterium

Pred. No. is the number of results predicted by chance to have a

```
JOURNAL    Patent: WO 0231158-A 1 18-APR-2002;
           ARCHER-DANIELS MIDLAND COMPANY (US)
FEATURES             Location/Qualifiers
    source           1..3474
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
    CDS              1..3474
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAD38136.1"
                     /db_xref="GI:21712842"
                     /translation="MTAITLGGLLLKGIITLVSTHTSSTLPAFKKILVANRGEIAVRA
                     FRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYLDIDELIGAKKVK
                     ADAIYPGVGFLSENAQLARECAENGITPIGPTPEVLLDTGDKSRAVTAAKKAGLPVLA
                     ESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAA
                     FGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDP
                     ELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVD
                     LVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGITAYRSPG
                     GAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIG
                     FLRALLREEDFTSKRIATGFPIGDHPHLLQAPPADDEQGRIIDYLADVTVNKPHGVRPK
                     DVAAPIDKLPNIKDLPLPRGSRRDLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLL
                     ATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELRREAMPN
                     VNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVL
                     ETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAA
                     VTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSL
                     SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFBESGTPGPTGRVYRHEIPGGQ
                     LSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVAGVDP
                     ADFAADPQKYDIPDSVIAFGRELGIGNPPGGWPEPLRTRALEGRSEGKAPUTEVPEEEQ
                     AHLDADDSKPRRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIR
                     LPDVRTPLLVRLDAISEPDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSSN
                     KGHVAAPFAGVVTVTVAEGDEVKAGDVAIIEAMKMEATITASVDGKIERVVPAATK
                     VEGGDLIVVVS"
BASE COUNT    749 a     1013 c     950 g     762 t
ORIGIN

Query Match        100.0%;  Score 3474;  DB 6;  Length 3474;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 3474;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy    1  GTGCTGCTATCACCCCTGGGCGGTCTTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GTGCTGCTATCACCCCTGGGCGGTCTTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT  60

Qy   61  CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC  120

Qy  121  GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGCCGTAGCTATTTACCCC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGCCGTAGCTATTTACCCC  180

Qy  181  CGTGAAGATCGGGGATCATTCCACCCGTCTTTTGCTTCTGACATCGATGAAATTATCGGTACT  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CGTGAAGATCGGGGATCATTCCACCCGTCTTTTGCTTCTGACATCGATGAAATTATCGGTACT  240

Qy  241  GAAGGCTCAAGGCGTATTACCCGGGATATGGCTTCCTGCTGCAGCTAAAAA  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  GAAGGCTCAAGGCGTATTACCCGGGATATGGCTTCCTGCTGCAGCTAAAAA  300

Qy  301  GTTAAAGCAGATGCTGCGGCGAAAACGGTCATTTATTGGCCCAGAGGTTCCTGATCTC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GTTAAAGCAGATGCTGCGGCGAAAACGGTCATTTATTGGCCCAGAGGTTCCTGATCTC  360

Qy  361  CGGCAGTCGGCGAAAACGGGCATTATTCACCCGGGGATAATGGCTTCCTGAAAATGCCC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CGGCAGTCGGCGAAAACGGGCATTATTCACCCGGGGATAATGGCTTCCTGAAAATGCCC  420

Qy  421  ACCGGTGATAAGTCTCGTGCGCGTAACCGGCCCGCGGAAGGCTGGTCTGTCGTCTGGCG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  ACCGGTGATAAGTCTCGTGCGCGTAACCGGCCCGCGGAAGGCTGGTCTGTCGTCTGGCG  480

Qy  481  GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGGCTGAAGGCCAGACTTAC  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGGCTGAAGGCCAGACTTAC  540

Qy  541  CCCATCTTTGTAAAGGCAGTGCCGGTGCGGCGACGCGGTATGCGCTTGTTTCTTCA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CCCATCTTTGTAAAGGCAGTGCCGGTGCGGCGACGCGGTATGCGCTTGTTTCTTCA  600

Qy  601  CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC  660

Qy  661  GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC  720

Qy  721  CTTGGCGATGCCACTGGAGAAGTGTACACCTTTATGAACGTGACTGCTCCACTGCGAGCGT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  CTTGGCGATGCCACTGGAGAAGTGTACACCTTTATGAACGTGACTGCTCCACTGCGAGCGT  780

Qy  781  CGTCACCAAAAAGTTGTCCAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  CGTCACCAAAAAGTTGTCCAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840

Qy  841  CGCATTGTGCGGATGCAGTAAAGTTCTGCGCGCTCCATTGGTTACCAGGGCGCGGGAACC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CGCATTGTGCGGATGCAGTAAAGTTCTGCGCGCTCCATTGGTTACCAGGGCGCGGGAACC  900

Qy  901  GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC  960

Qy  961  CAGGTTGAGCAGACACCGTGACTGGACCGAGGTGGACCTGGTGGTGAAGGCGCAGATG  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  CAGGTTGAGCAGACACCGTGACTGGACCGAGGTGGACCTGGTGGTGAAGGCGCAGATG  1020

Qy 1021  CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGCAAGACC  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGCAAGACC  1080

Qy 1081  CACGGTGCAGCACTGCCAGTGCCGCGATCACCACCGGAGATCCAAACAACGGCTTCCGCCCA  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  CACGGTGCAGCACTGCCAGTGCCGCGATCACCACCGGAGATCCAAACAACGGCTTCCGCCCA  1140

Qy 1141  GATACCGGAACTATCACCCGCGTACCCGTCACCAGGCGGAGCTGGCGTTCGACGGT  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  GATACCGGAACTATCACCCGCGTACCCGTCACCAGGCGGAGCTGGCGTTCGACGGT  1200

Qy 1201  GCAGCTCAGCTCGGTGGCGAAATCACCGGTAGCCACGGTAGCTATTTGACT  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  GCAGCTCAGCTCGGTGGCGAAATCACCGGTAGCCACGGTAGCTATTTGACT  1260

Qy 1261  TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCGTGTGCGCGTTGCTGCGGGAAGAGGAC  1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCGTGTGCGCGTTGCTGCGGGAAGAGGAC  1320

Qy 1321  ACCGTGTCTGGGTGTGCAACCAACATTGGTTTCTTGTTCTTGCGGATCACCCACCTCCTTCAGGCT  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  ACCGTGTCTGGGTGTGCAACCAACATTGGTTTCTTGTTCTTGCGGATCACCCACCTCCTTCAGGCT  1380

Qy 1381  TTCACTTCCAAGCGCATCGCCACCGGAGACGCATCCTGGATTACTTGGCAGAGTCGCAACATCGATAAGCTGCACCAATCGATAAGCTGCACCAAATCGCGTGAAC  1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  TTCACTTCCAAGCGCATCGCCACCGGAGACGCATCCTGGATTACTTGGCAGAGTCGCAACATCGATAAGCTGCACCAATCGATAAGCTGCACCAAATCGCGTGAAC  1500

Qy 1441  CAACCTGCGGATGATGAGCAGCAGGACGGACGCATCCTGGATTACTTGGCAGAGTCGCTGCCGATGTCCAACATC  1500
Db 1441  CAACCTGCGGATGATGAGCAGCAGGACGGACGCATCCTGGATTACTTGGCAGAGTCGCTGCCGATGTCCAACATC  1500

Qy 1501  AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCAACATC  1560
Db 1501  AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCAACATC  1560

Qy 1561  AAGGAATCTGCCACTGCCACGGCGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCTGAGACTTAC  1620
```

```
Db  1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCAGCCGCG 1620

Qy  1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680

Qy  1681 GCACACCAGTCTTTGCTTGCTGGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 GCACACCAGTCTTTGCTTGCTGGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740

Qy  1741 GCCGTCGCAAAGCTGACTCCTGAGCTTGTCCGTGGAGGCGTGGGGCGGCGGCCGGACCTAC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GCCGTCGCAAAGCTGACTCCTGAGCTTGTCCGTGGAGGCGTGGGGCGGCGGCCGGACCTAC 1800

Qy  1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGCGAGCTGCCCGAG 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGCGAGCTGCCCGAG 1860

Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTTGGGATACACC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTTGGGATACACC 1920

Qy  1921 CCGTACCCAGAGCTCCGTCTGCCCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCCGTGGAC 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CCGTACCCAGAGCTCCGTCTGCCCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCCGTGGAC 1980

Qy  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGGTCGTTCAGCAATCGACGCA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGGTCGTTCAGCAATCGACGCA 2040

Qy  2041 GTCCTGGAGACCAACACCCGGGTAGCCGGGTGGGGGGTAGCCGAGTGGTATGCGTATTCTGTGATCTCTCT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTCCTGGAGACCAACACCCGGGTAGCCGGGTGGGGGGTAGCCGAGTGGTATGCGTATTCTGTGATCTCTCT 2100

Qy  2101 GATCCAAATGAAAAGCTCTACACACCTGGATTACTACACCCTGGATTACTACACCCTAAAGATGGCAGGAGATCGTC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 GATCCAAATGAAAAGCTCTACACACCTGGATTACTACACCCTGGATTACTACACCCTAAAGATGGCAGGAGATCGTC 2160

Qy  2161 AAGTCTGGCGCTCACATTCTGCCGTCTGATATGCGGCCGGAATTCGATCTGCCAGTGCACGTGCACACC 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AAGTCTGGCGCTCACATTCTGCCGTCTGATATGCGGCCGGAATTCGATCTGCCAGTGCACGTGCACACC 2220

Qy  2221 GTAACCAAGCTGGTCACCGGACTGGCGCGCCGTGAATTCGATCTGCCCAGTGCACGTGCACACC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 GTAACCAAGCTGGTCACCGGACTGGCGCGCCGTGAATTCGATCTGCCCAGTGCACGTGCACACC 2280

Qy  2281 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTTGCTGCGGCTTGCTGCGCAGCTTCAAGGCTGGTCAGAT 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTTGCTGCGGCTTGCTGCGCAGCTTCAAGGCTGGTCAGAT 2340

Qy  2341 GCTGTTGACGGTGCTTCCGCCACTGTCTGCGCCACCACTGTCTGCGCCATCCCTGTCGCC 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GCTGTTGACGGTGCTTCCGCCACTGTCTGCGCCACCACTGTCTGCGCCATCCCTGTCGCC 2400

Qy  2401 ATTGTTGCTGCATTCGCATCGGCGGCGATACCCGGTTTGAGGCTTCGAGGGCTGTTTCT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 ATTGTTGCTGCATTCGCATCGGCGGCGATACCCGGTTTGAGGCTTCGAGGGCTGTTTCT 2460

Qy  2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTGAGTCTGCGAACC 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTGAGTCTGCGAACC 2520

Qy  2521 CCAGGCCCAACCGGTGCCGTCGTACCGCCCACGGCGGACAGTTGTCCAACCTG 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 CCAGGCCCAACCGGTGCCGTCGTACCGCCCACGGCGGACAGTTGTCCAACCTG 2580

Qy  2581 CGTGCACAGGCCACCGGCCACTGGGCCTTGCTGATCGCTTCGAGCTCGTGATCGCTTCGAGCTCGTGAAGACAACTAC 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 CGTGCACAGGCCACCGGCCACTGGGCCTTGCTGATCGCTTCGAGCTCGTGATCGCTTCGAGCTCGTGAAGACAACTAC 2640

Qy  2641 GCAGCCGTTAATGAGATGCTGGGACGCCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 GCAGCCGTTAATGAGATGCTGGGACGCCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
```

```
Qy  2701 GGCGACCTGCCACTGCCACTGCCACTTGGTTGCGGTGCGGGTGTGAGATCCAGCAGACTTTGCTGCAGAC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GGCGACCTGCCACTGCCACTGCCACTTGGTTGCGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC 2760

Qy  2761 CCACAAAGTACGACATCCCAGACTCGTCATCGCGGCGGAGCTTGGTAAC 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 CCACAAAGTACGACATCCCAGACTCGTCATCGCGGCGGAGCTTGGTAAC 2820

Qy  2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCACTGCGCACCGCCGCACCTGGAAGGCCGCTCCGAGGGC 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCACTGCGCACCGCCGCACCTGGAAGGCCGCTCCGAGGGC 2880

Qy  2881 AAGGCACCTCTGACGGAAGTTCCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGTGATGATTCC 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 AAGGCACCTCTGACGGAAGTTCCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGTGATGATTCC 2940

Qy  2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCGTTCCGAAGCCAACCAGAAGAGTTCCTC 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCGTTCCGAAGCCAACCAGAAGAGTTCCTC 3000

Qy  3001 GAGCACCGTGCGCCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTCTTCTACGGA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 GAGCACCGTGCGCCGCTTCGGCAACACCTCTGCGCTGATGATCGTGAATTCTTCTACGGA 3060

Qy  3061 CTGGTCGAGGGCCGCGGACTTGATCCGCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 CTGGTCGAGGGCCGCGGACTTGATCCGCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120

Qy  3121 CGCCTGATGGCATCTTCGAGCCAGACGATAAGGTATGCGGCAATGTTGTGGCCAAGCGTC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CGCCTGATGGCATCTTCGAGCCAGACGATAAGGTATGCGGCAATGTTGTGGCCAAGCGTC 3180

Qy  3181 AACGGCCAGATCCGCCCAATGCGTGCGTCCGTTGAGTCTCGTGACCGCGCCAACC 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 AACGGCCAGATCCGCCCAATGCGTGCGTCCGTTGAGTCTCGTGACCGCGCCAACC 3240

Qy  3241 GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCTCGGTGTGTC 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCTCGGTGTGTC 3300

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTGCAATCATCGAG 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTGCAATCATCGAG 3360

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGACGGCAAGATTGAACGCGTTGTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGACGGCAAGATTGAACGCGTTGTG 3420

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGTTGGCGACTTGATCGTCGTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTTCCTGCTGCAACGAAGGTGGAAGTTGGCGACTTGATCGTCGTTCCTAA 3474
```

```
RESULT 2
AX453603
LOCUS       AX453603              3474 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent WO0231158.
ACCESSION   AX453603
VERSION     AX453603.1  GI:21712843
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1
  AUTHORS   Hanke,P.D.
  TITLE     Feedback-resistant pyruvate carboxylase gene from corynebacterium
  JOURNAL   Patent: WO 0231158-A 3 18-APR-2002;
            ARCHER-DANIELS-MIDLAND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..3474
```

```
ORIGIN
                    /organism="Corynebacterium glutamicum"
                    /db_xref="taxon:1718"
BASE COUNT      749 a    1013 c     950 g     762 t

Query Match           100.0%;    Score 3474;    DB 6;    Length 3474;
Best Local Similarity 100.0%;    Pred. No.0;
Matches 3474;    Conservative 0;    Mismatches 0;    Indels 0;    Gaps 0;

Qy      1 GTGACTGCTATCACCCTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACTGCTATCACCCTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy     61 CACCACATCTTCAACGCTTCCAGCACTTCAAAAAGATCTTCGTAGCAAACCGGCGGAAATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CACCACATCTTCAACGCTTCCAGCACTTCAAAAAGATCTTCGTAGCAAACCGGCGGAAATC 120

Qy    121 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCTATTTACCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCTATTTACCCC 180

Qy    181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCCGAAGCTGTCCGCATTGGTACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCCGAAGCTGTCCGCATTGGTACT 240

Qy    241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300

Qy    301 GTTAAAGCAGAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTTAAAGCAGAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy    361 CGCGAGTGCGCGGAAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCGAGTGCGCGGAAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420

Qy    421 ACCGGTGATAAGTCTCGTCGCGGTAACCGCGCGGAGAGGCTGTCTGCCAGTTTTGGGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACCGGTGATAAGTCTCGTCGCGGTAACCGCGCGGAGAGGCTGTCTGCCAGTTTTGGGCG 480

Qy    481 GAATCCCACCCCGAGCAAAAACATCGATGACATCGTTAAAGCGCTGAAAGCGCCAGACTTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAATCCCACCCCGAGCAAAAACATCGATGACATCGTTAAAGCGCTGAAAGCGCCAGACTTAC 540

Qy    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTCGGCGACGCGGTATGCGGCTTTGTTTCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTCGGCGACGCGGTATGCGGCTTTGTTTCTTCA 600

Qy    601 CCTGATGAGCTCGCAAATTGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTGATGAGCTCGCAAATTGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660

Qy    661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCGAGATC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCGAGATC 720

Qy    721 CTTGGCGGATCGCACTGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTTGGCGGATCGCACTGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780

Qy    781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCAGCATTTGGATCCGAAACTGCGTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCAGCATTTGGATCCGAAACTGCGTGAT 840

Qy    841 CGCATTTGTGCCGATGCAGTAAAGTTCTGCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGCATTTGTGCCGATGCAGTAAAGTTCTGCGCTCCATTGGTTACCAGGGCGCGGGAACC 900

Qy    901 GTGGAATTCTTGGTCGATGAGGCAACCACCAGTTTCATCGAAATGAACCCACGTATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GTGGAATTCTTGGTCGATGAGGCAACCACCAGTTTCATCGAAATGAACCCACGTATC 960

Qy    961 CAGGTTGAGCAGCACCGTGACTGCTGAAGAGTCACCGAGGTGGACCTGGTGAGGCGCAGATG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CAGGTTGAGCAGCACCGTGACTGCTGAAGAGTCACCGAGGTGGACCTGGTGAGGCGCAGATG 1020

Qy   1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGCTGCTGACCCAAGATAAGATCAAGACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGCTGCTGACCCAAGATAAGATCAAGACC 1080

Qy   1081 CACGGTGCAGCACTGCAGTGCCGATCACACGAAGATCCAAACAACAGCGGCTTCCGCCCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CACGGTGCAGCACTGCAGTGCCGATCACACGAAGATCCAAACAACAGCGGCTTCCGCCCA 1140

Qy   1141 GATACCGGAACTATCACCCGCGTACCGCGTCACCAGGCGGAGCTGGCCTTCGTCTTGACGGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GATACCGGAACTATCACCCGCGTACCGCGTCACCAGGCGGAGCTGGCCTTCGTCTTGACGGT 1200

Qy   1201 GCAGCTCAGCTCGGTGGGGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCAGCTCAGCTCGGTGGGGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260

Qy   1261 TGCCGTGGTTCCGACTTTGAAACTGTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCGTGGTTCCGACTTTGAAACTGTCGTGCACAGCGCGCGTTGGCTGAGTTC 1320

Qy   1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCGCGGAAGAGGAC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCGCGGAAGAGGAC 1380

Qy   1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCACCTCCTTCAGGCT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCACCTCCTTCAGGCT 1440

Qy   1441 CCACCTTGCGGATGATGAGCAGCAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCACCTTGCGGATGATGAGCAGCAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500

Qy   1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGATAAGCTCCCAACATC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGATAAGCTCCCAACATC 1560

Qy   1561 AAGGATCTGCCACTGCCACGGTTCCCGTGAAGCAGCTTGGCCCAGCGGCG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AAGGATCTGCCACTGCCACGGTTCCCGTGAAGCAGCTTGGCCCAGCGGCG 1620

Qy   1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680

Qy   1681 GCACACCAGTCTTTGCTTGCTGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCACACCAGTCTTTGCTTGCTGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 1740

Qy   1741 GCCGTCGCAAAGCTGACTCCTGAGCGTTGTCCGTGGAGGCGGCGGCGGCGACCTAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GCCGTCGCAAAGCTGACTCCTGAGCGTTGTCCGTGGAGGCGGCGGCGGCGACCTAC 1800

Qy   1801 GATGTGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGGCGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GATGTGCGATGCGTTTCTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTGCGGCGAG 1860

Qy   1861 GCGATGCGAATGTAAACATTCAGATGCTGCTTCGCGCTTCGCGGCCCAACACCGTGGGATACACC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCGATGCGAATGTAAACATTCAGATGCTGCTTCGCGCTTCGCGGCCCAACACCGTGGGATACACC 1920

Qy   1921 CCGTACCCAGACTCCGTCTGCCGTCCAGCTGCCAGCAGCTGCCAGCTCCCGGCGTGGGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCGTACCCAGACTCCGTCTGCCGTCCAGCTGCCAGCAGCTGCCAGCTCCCGGCGTGGGAC 1980

Qy   1981 ATCTTCCGCCATCTTCGACGCGGCTTAACGACGTCTCCCAGCATGCGTCAGCAATCGACGCA 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATCTTCCGCCATCTTCGACGCGGCTTAACGACGTCTCCCAGCATGCGTCAGCAATCGACGCA 2040

Qy   2041 GTCCTGGAGACCAACACCCGGCGTAGCCGAGGTGCGTATGGCTTATTCGTGGTGATCTCTCT 2100
```

Left column:

```
Db   2041  ||||||||||||||||||||||||||||||||||||||||||||  2100
           GTCCTCGGAGACCAACACCGCGCGGTAGCCGAGGTGGCTATGCGGCTTATTCGTGATCTCT

Qy   2101  GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC  2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC  2160

Qy   2161  AAGTCTGGCGCGTCACATTCTGGCCACTCTGCGCACGCTGCTGATCTGCCAGTGCACACC  2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  AAGTCTGGCGCGTCACATTCTGGCCACTCTGCGCACGCTGCTGATCTGCCAGTGCACACC  2220

Qy   2221  GTAACCAAGTCGGTCACGGGGTGCCAGTGCTACCTGCACCACTCTGCGCAGCTCAAGCTGGTGCAGCT  2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  GTAACCAAGTCGGTCACGGGGTGCCAGTGCTACCTGCACCACTCTGCGCAGCTCAAGCTGGTGCAGCT  2280

Qy   2281  CACGACACTGCGGGTGCCAGTGGCCGCACCACTCTGCGCAGCTCAAGCTGGTGCGAGAT  2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  CACGACACTGCGGGTGCCAGTGGCCGCACCACTCTGCGCAGCTCAAGCTGGTGCGAGAT  2340

Qy   2341  GCTGTTGACGGTGCTTCCGCACCACTCTGCGCACCACCTCCCAGCCATCCCTGTCGCC  2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  GCTGTTGACGGTGCTTCCGCACCACTCTGCGCACCACCTCCCAGCCATCCCTGTCGCC  2400

Qy   2401  ATTGTTGCTGCATTCGCGCACACCCGTTGAGCCTCGAGGCTGTTTCT  2460
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  ATTGTTGCTGCATTCGCGCACACCCGTTGAGCCTCGAGGCTGTTTCT  2460

Qy   2461  GACCTCGAGCGCTACTGGGAAGCTGTGCGCGGACTGTGCCCATTTGAGTCTGGAACC  2520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  GACCTCGAGCGCTACTGGGAAGCTGTGCGCGGACTGTGCCCATTTGAGTCTGGAACC  2520

Qy   2521  CCAGGGCCCAACCGGTCGCGGTCTACCGCCACGAAATCCCAGGCGGACAGTGTCCAACCTG  2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  CCAGGGCCCAACCGGTCGCGGTCTACCGCCACGAAATCCCAGGCGGACAGTGTCCAACCTG  2580

Qy   2581  CGTGCACAGGCCACCGCACCTGGGCCTTCGATCGCTTGCTGAGCTCATCATCGAAGACAACTAC  2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  CGTGCACAGGCCACCGCACCTGGGCCTTCGATCGCTTGCTGAGCTCATCATCGAAGACAACTAC  2640

Qy   2641  GCAGCCGTTAATGATGAGTGCTGGACGCGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  GCAGCCGTTAATGATGAGTGCTGGACGCGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2700

Qy   2701  GGGGACCTCGCACTCCACCTCGGTTGGTGCGGGGTGTAGATCCAGCAGAGACTTTGCTGCAGAC  2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  GGGGACCTCGCACTCCACCTCGGTTGGTGCGGGGTGTAGATCCAGCAGAGACTTTGCTGCAGAC  2760

Qy   2761  CCACAAAGTACGACATCTGTCATCGCGTTCCTGCGCGCCGAGCTTGTAAC  2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  CCACAAAGTACGACATCTGTCATCGCGTTCCTGCGCGCCGAGCTTGTAAC  2820

Qy   2821  CCTCCAGGTGCTGGCCAGAACCACTGCGCACGCACCGCGCGAAGGCCGCTCCGAAGGC  2880
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821  CCTCCAGGTGCTGGCCAGAACCACTGCGCACGCACCGCGCGAAGGCCGCTCCGAAGGC  2880

Qy   2881  AAGGACCTCTGACGACGAAGTTCCTGAGGAAGAGCAGCAGGCGCACTCGACGCTGATGATTCC  2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881  AAGGACCTCTGACGACGAAGTTCCTGAGGAAGAGCAGCAGGCGCACTCGACGCTGATGATTCC  2940

Qy   2941  AAGGACGTCGCAACAGCTCGCTTCAACCGCCTCAACCGGAAGAGCCAACCGGAGAGTTCCTC  3000
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941  AAGGACGTCGCAACAGCTCGCTTCAACCGCCTCAACCGGAAGAGCCAACCGGAGAGTTCCTC  3000

Qy   3001  GAGCACCGGTGCCGCGCGTTCGGCAACAACCTCTGCGCGCTGATGATCGTGAATTCTTCACGGA  3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001  GAGCACCGGTGCCGCGCGTTCGGCAACAACCTCTGCGCGCTGATGATCGTGAATTCTTCACGGA  3060

Qy   3061  CTGGTTCGAGGGCCGCGAGAGTTTGATCCGGTCCTGCCAGAGATGTGCCGACCCCACTGCTTGTT  3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061  CTGGTTCGAGGGCCGCGAGAGTTTGATCCGGTCCTGCCAGAGATGTGCCGACCCCACTGCTTGTT  3120

Qy   3121  CGCCTGATGCGGATCTCTGAGCCAGACGCGATAAGGGTATGCGGCAATGTTGTGCGGCCAACGTC  3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column:

```
Db   3121  CGCCTGATGCGGATCTCTGAGCCAGACGCGATAAGGGTATGCGGCAATGTTGTGCGGCCAACGTC  3180

Qy   3181  AACGGCCAGATCCGCCCATGCGTGCGTGCGACCAGATCGTC  3240
           |||||||||||||||||||||||||||||||||||||||||
Db   3181  AACGGCCAGATCCGCCCATGCGTGCGTGCGACCAGATCGTC  3240

Qy   3241  GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCTGACCATTCGCTGGTGTTGTC  3300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241  GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCTGACCATTCGCTGGTGTTGTC  3300

Qy   3301  ACTGTGACTGTTGCTGAAGGTAGGTGATGGAGGTCAAGGCTCGAGATGCGAGTCGCAATCATCGAG  3360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301  ACTGTGACTGTTGCTGAAGGTAGGTGATGGAGGTCAAGGCTCGAGATGCGAGTCGCAATCATCGAG  3360

Qy   3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGTTGACGGCAAGATTGAACGCGTTGTG  3420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGTTGACGGCAAGATTGAACGCGTTGTG  3420

Qy   3421  GTTCCTGCTGCAACGAAGGTGGCGACTTGATCGTCGTCGTTCCTAA  3474
           |||||||||||||||||||||||||||||||||||||||||||||
Db   3421  GTTCCTGCTGCAACGAAGGTGGCGACTTGATCGTCGTCGTTCCTAA  3474
```

RESULT 3
AR123974

```
LOCUS       AR123974        Sequence 1  from patent US 6171833.          3621 bp    DNA        linear     PAT 16-MAY-2001
DEFINITION  AR123974        Sequence 1 from patent US 6171833.
ACCESSION   AR123974
VERSION     AR123974.1  GI:14109335
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3621)
  AUTHORS   Sinskey,A.J., Lessard,P.A. and Willis,L.B.
  TITLE     Pyruvate carboxylase from corynebacterium glutamicum
  JOURNAL   Patent: US 6171833-A 1 09-JAN-2001;
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="unknown"
BASE COUNT      782 a   1037 c   1002 g    800 t
ORIGIN
```

Query Match   97.8%; Score 3398.8; DB 6; Length 3621;
Best Local Similarity   98.6%; Pred. No. 0;
Matches 3427; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy      1  GTGACTGCTATCACCCTGGGCGTCTTCGTTGTTGAAAGGAATAATTACTCTAGTGTCGACT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148  GTGACTGCTATCACCCTGGGCGTCTTCGTTGTTGAAAGGAATAATTACTCTAGTGTCGACT  207

Qy     61  CACACATCTTCAACGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGGGCGGCGAAATC  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208  CACACATCTTCAACGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGGGCGGCGAAATC  267

Qy    121  GCGGTCCGTGTTCCGTGCAGCACTCGAACTCGAACTCGAAACCGGTGCAGCGTAGCTATTACCCC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268  GCGGTCCGTGTTCCGTGCAGCACTCGAACTCGAACTCGAAACCGGTGCAGCGTAGCTATTACCCC  327

Qy    181  CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTTTGCTTCCGAAGCTGTCCGCATTGGTACT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328  CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTTTGCTTCCGAAGCTGTCCGCATTGGTACC  387

Qy    241  GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388  GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA  447

Qy    301  GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448  GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC  507

Qy    361  CGCGAGTGCGCGGAAAACGCATTATTTTATTGGCCAACCCCAGAGGTTCTTGATCTC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    508  CGCGAGTGTGCCGAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC  567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501  AAGCCTCATGTGTGGCGTCCAAAGAGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC  1560

Qy    421  ACCGGTGATAAGTCTCGTGCGCGGTAACCGCGAAGAGGCTGGTCTGTCCAGTTTTTGGCG  480
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648  AAGCCTCATGTGTGGCGTCCAAAGAGATGTTGCAGCACCTCCTATCGATAAGCTGCCTAACATC  1707

Db    568  ACCGGTGATAAGTCTCGCGCGGTAACCGCGAAGAGGCTGGTCTCCCAGTTTTGGCG  627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1561  AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGGTGAAGCAGCTTGCCCCAGCCGCG  1620

Qy    481  GAATCCACCCCAGCCAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708  AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGGTGAAGCAGCTTGCCCCAGCCGCG  1767

Db    628  GAATCCCGAGCAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGGCCAGACTTAC  687
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621  TTTGCTCGTGATCTCCGTGAGCAGGAGGACCACTGGCAGTTACTGATACCACCTTCCGCGAT  1680

Qy    541  CCATCTTGTAAAGGCAGTTGCCCGGTGGTGGCCGACCGGCGGTATGCGGTTTCTTCA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688  CCCATCTTTGTGAAGGCAGTTGCCCGGTGGCGGCGTATGCGGTTTTGTTGCTTCA  747

Db    601  CTTGATGAGCTCCGCAAATTGCAACAGAGACATCTCGTGAAGCTGAAGCGGCATTCGGC  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681  GCACACCAGTCTTTGCTTGCGACCCGACCCGAGTCCGCTCATTCGCACTGAAGCGCGCGGCAGAG  1740

Qy    661  GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748  CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC  807

Db    668  GACGGTTCGGTATATGTCGAACGTGTGTGATTAACCCTCAGCATATTGAAGTGCAGATC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1741  GCCGTGCAAAGCTGACTCCTGAGCTTTTGTCCGTGAGGCTTTTGTCCGTGAGGCGCGGGGGCGTGGGCGCGGGACCTAC  1800

Qy    721  CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808  GATGGCGGGGTATATGTCGAACGTGCTGTGTGTGATTAACCCTCAGCATATTGAAGTGCAGATC  867

Db    721  CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1801  GATGTGGCGATGGTGTTCCTCTTGAGATCCGTGGGACATCCGTGGGACAGGCTCGACAGGCTCGACGAGCTGCCGCGAG  1860

Qy    781  CGTCACCAAAAGTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868  CTTGGCGATCACCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  927

Db    928  CGTCACCAAAAGTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  987
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1861  GCGATGCCGAATGTAAAACATTCAGATGCTGCTTCGCCGGCGCCCAACACCGTGGGATACACC  1920

Qy    841  CGCATTTGTGCGGATGCAGTAGTAAGTTCTGCCCGCTCCATTGGTTACCAGGGGCGCGGGAACC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    988  CGTACCCAGACTCCGTCTGCCCGCGCTGCTGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC  1980

Db    988  GTGGAATTCTTCGGTCGATGAAAAGGGCAACCACCGTCTTCATCGAAATGAACCCACGTATC  1047
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1921  CCGTACCCAGACTCCGTCTGCCCGCGCTGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC  1980

Qy    901  GTGGAATTCTTCGGTCGATGAAAAGGGCAACCACCGTTTTCATCGAAATGAACCCACGTATC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2068  CCGTACCCAGACTCCGTCTGCCCGCGCTGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC  2127

Db   1048  GTGGAATTCTTCGGTCGATGAAAAGGGCAACCACCGTCTTCATCGAAATGAACCCACGTATC  1107
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1981  ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGATCGTCTTATCTCGTGATCTCTCT  2040

Qy    961  CAGGTTGAGCACACCGTTGACTGTGACCTTGACCGAGGTCGACCTTGGTCGAAGGGGCAGATG  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2128  ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGATCGTCTTATCTCGTGATCTCTCT  2187

Db   1108  CAGGTTGAGCACACCGTTGACTGACTGAAGAAGTCACCGGAGGTCGGCAGATG  1167
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041  GTCCTGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTAAAGATGCGCAGAGGAGATCGTC  2100

Qy   1021  CGCTTGGCTGCTGGTGCAACCGTCAGTGCGCCGCATCACCACGGAAGAGTCAAACACGGCTTCCGCCCA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2188  GTCCTGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTAAAGATGCGCAGAGGAGATCGTC  2247

Db   1168  CAGGTTGGCTGCTGGTGCAACCGTCAGTGCGCCGCATCACCACGGAAGAGTCAAACACGGCTTCCGCCCA  1227
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2101  GATCCAAATGAAAAGCTCTACACCCTCGATATTACTACCTAAAGATGCGAGGAGATCGTC  2160

Qy   1081  CACGGTGCGAGCACTGCGACTGCCCGCATCACCACGGAAGAATCCAAACACACGGCTTCCGCCCCA  1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2248  GATCCAAATGAAAAGCTCTACACCCTCGATATTACTACCTAAAGATGCGAGGAGATCGTC  2307

Db   1228  CACGGTGCGAGCACTGCGAGCTGCCCGCATCACCACGGAAGAATCAAACACACGGCTTCCGCCCA  1287
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2161  AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCGTGCTGCTTCGCCCAGCTGCG  2220

Qy   1141  GATACCGGAACTATCACCGCGTACCGGTCACCGCTCACCGGCGGAGCTGGTTCGTCGTTGTGACGGT  1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2308  AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCGTGCTGCTTCGCCCAGCTGCG  2367

Db   1288  GATACCGGAACTATCACCGCGTACCGGTCACCGGCGAGCTGGTTCGTCGTTGTGACGGT  1347
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2221  GTAACCAAGCTGGTCACCGCACTGCGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC  2280

Qy   1201  GCAGCTCAGCTCGGTGGCGAAATCACCGGCGACACTTGACTCGTGTGAAAATGACC  1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2368  GTAACCAAGCTGGTCACCGCACTGCGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC  2427

Db   1348  GCAGCTCAGCTCGGTGGCGAAATCACCGGCGACACTTGACTCGTGTGAAAATGACC  1407
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2281  CACGACACTGCGGGTGCCAGTTGCCGTACCTACTTGCTCGAGCTCGGTGCAGCTGGTGCAGAT  2340

Qy   1261  TGCCGGTCGTTCCGACTTTGAAACTGCGTTCTGCTCCGTGCACAGCCGCGCGTTGGCTGAGTTC  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408  TGCCGGTCGTTCCGACTTTGAAACTGCGTTCTGCTCCGTGCACAGCCGCGCGTTGGCTGAGTTC  1467

Db   1408  CACGACACTGCGGGTGCCAGTTGCCGTACCTACTTGCTCGCAGCTCGGTGCAGCTGGTGCAGAT  1487
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2341  GCTGTTGACGGTGCTTCCAGCAACACTGTCTGGCGACCACTCCAGCAGCCATCCGTCTGCC  2400

Qy   1321  ACGGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTCGGGCGTTGCTGCGGGGAAGAGGAC  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468  ACCGTGTCTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTCGGGCGTTGCTGCGGGGAAGAGGAC  1527

Db   1488  GCTGTTGACGGTGCTTCCAGCAACACTGTCTGGCGACCACTCCAGCAGCCATCCGTCTGCC  1547
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2401  ATTGTTGCTGCGATTGCAACACCAACCGGTTTGAGCGTTTGAGCCTGGAGGCTGTTTCT  2460

Qy   1381  TTCACTTCCAAGCGCATCGCACCGGATTTATCGGCGCATTGCCGATGGCCTTCCTTCAGGGCT  1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528  TTCACTTCCAAGCGCATCGCACCGGATTTATCGGCGCATTGCCGATGGCCTTCCTTCAGGGCT  1587

Db   1528  ATTGTTGCTGCGATTGCAACACCAACCGGTTTGAGCGTTTGAGCCTGGAGGCTGTTTCT  2607
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2461  GACCTGCGAGCCGTACTGGGAAGCTGTGCGGACTGTGTACCTGCCATTTGAGTCTGGAACC  2520

Qy   1441  CCACCTGCCGGATGAGCAGGAGGACCGGCAGCGAGGACCGCATCCTGGATTACTTGCACCGTGTCACCGTGAAC  1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2608  GACCTGCGAGCCGTACTGGGAAGCTGTGCGGACTGTGTACCTGCCATTTGAGTCTGGAACC  2667

Db   1588  CCACCTGCTGATGATGAGCAGCAGGAGGACGCAGGGACGCGATCCTGGATTACTTGCACCGTGAAC  1647
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2521  CCAGGCCCAACCGGTCGCGTCTACCGCCGACAAATCCGGAGCAGAGTTGTCCAACCTG  2580

           CCAGGCCCAACCGGTCGCGTCTACCGCCGACAAATCCGGAGCAGAGTTGTCCAACCTG  2580
Db   2668  CCAGGCCCAACCGGTCGCGTCTACCGCCGACAAATCCGGAGCAGAGTTGTCCAACCTG  2727
```

```
Qy  2581  CGTGCACAGGCCACCCGCACTGGGCCTTGCTGTATCGCTTCGAGCTCATCGAAGACAACTAC  2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2728  CGTGCACAGGCCACCCGCACTGGGCCTTGCTGTATCGCTTCGAGCTCATCGAAGACAACTAC  2787

Qy  2641  GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2788  GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2847

Qy  2701  GGGACCTCGCACTCCACTCGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCGAGAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2848  GGGACCTCGCACTCCACTCGTTGGTGCGGGTGTGATCCAGCAGACTTTGCTGCGAGAT  2907

Qy  2761  CCACAAAAGTACGACATCCCAGACTCGTCATCGGCGTTCCTGCGCGGCGAGCTTGGTAAC  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2908  CCACAAAAGTACGACACTCCCAGACTCGTCATCGCCGTTCCTGCGCGGCGAGCTTGGTAAC  2967

Qy  2821  CCTCCAGGTGCTGGCCAGAACCACTGCGCACCCGCGACGGGAAGGCCGCTCCGAAGGC  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2968  CCTCCAGGTGCTGCGCCAGAGCCACTGCGCACCCGCGACGGGAAGGCCGCTCCGAAGGC  3027

Qy  2881  AAGGACACCTCTGACGGAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC  2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3028  AAGGACACCTCTGACGGAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCC  3087

Qy  2941  AAGGAACGTGCAACAGCCTCAACCGCCTGCTGTTCCGGAAGCCAACCGAAGAGTTCCTC  3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3088  AAGGAACGTCCAATAGCCTCAACCGCCTGCTGTTCCGGAAGCCAACCGAAGAGTTCCTC  3147

Qy  3001  GAGCACCGTCGCCGCTTCGGCGAACACCTCTGCGCGTTCGGCGATCGTGAATTCTTCACGGA  3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3148  GAGCACCGTCGCCGCTTCGGCGAACACCTCTGCGCGTTCGGCGATCGTGAATTCTTCACGGC  3207

Qy  3061  CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGACCCCGCTGCTGTT  3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3208  CTGGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGACCCCGCTGCTGTT  3267

Qy  3121  CGCCTGGATCGCGATCTCTGAGCCAGAGGTATGCGCAATGTTGGCCAACGTC  3180
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  3268  CGCCTGGATCGCGATCTCTGAGCCAGAGGTATGCGCAATGTTGGCCAACGTC  3327

Qy  3181  AACGGCCAGATCCGCCCAATGCGTGCTGGCTACCGCTCCGTTGAGTCTGTCACCCGCAACC  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3328  AACGGCCAGATCCGCCCAATGCGTGCTGGCTACCGCTCCGTTGAGTCTGTCACCCGCAACC  3387

Qy  3241  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGACCCATTCGCTGGTGTGTC  3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3388  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGACCCATTCGCTGGTGTGTC  3447

Qy  3301  ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCGTGGAGATGCAGTCGCCAATCATCGAG  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3448  ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCGTGGAGATGCAGTCGCCAATCATCGAG  3507

Qy  3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTCTGTGACGCGGCAAGATTGAACGCGTTGTG  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3508  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTCTGTGACGCGGCAAAATCGATCGCGTTGTG  3567

Qy  3421  GTTCCTGCTGCAACGAAGGTGGAAGTTGGCGACTTGATCGTCGTCGTTCCTAA  3474
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3568  GTTCCTGCTGCAACGAAGGTGGAAGTTGGCGACTTGATCGTCGTCGTTCCTAA  3621
```

RESULT 4
AF038548

LOCUS   AF038548   3637 bp   DNA   linear   BCT 25-AUG-2000
DEFINITION   Corynebacterium glutamicum pyruvate carboxylase (pyc) gene,
             complete cds.
ACCESSION   AF038548
VERSION   AF038548.1   GI:2708716
KEYWORDS   .
SOURCE   Corynebacterium glutamicum.
ORGANISM   Corynebacterium glutamicum
             Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae;

REFERENCE   1 (bases 1 to 3637)
AUTHORS   Koffas,M.A., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and Stephanopoulos,G.
TITLE   Sequence of the Corynebacterium glutamicum pyruvate carboxylase gene
JOURNAL   Appl. Microbiol. Biotechnol. 50 (3), 346-352 (1998)
MEDLINE   99019028
PUBMED   9802220
REFERENCE   2 (bases 1 to 3637)
AUTHORS   Koffas,M.A.G., Ramamoorthi,R., Pine,W.A., Sinskey,A.J. and Stephanopoulos,G.
TITLE   Direct Submission
JOURNAL   Submitted (14-DEC-1997) Chemical Engineering, Massachusetts Insitute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

FEATURES   Location/Qualifiers
     source   1..3637
                     /organism="Corynebacterium glutamicum"
                     /strain="21253"
                     /db_xref="taxon:1718"
     gene   199..3621
                     /gene="pyc"
     CDS   199..3621
                     /gene="pyc"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="AAB92588.1"
                     /db_xref="GI:2708717"
                     /translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSPVKAYLDIEIIGAAKKVKADAIYPGYGFLSENAQL
                     ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQH
                     IEVQILGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI
                     GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEVTEVDLVKAQMRLAAGATLKEL
                     GLTQDKIKTHGAALQCRITEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
                     TGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPL
                     PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA
                     KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
                     YPDSVCRAFVKEAASGVDIFRIFDALNDVSQMRPAIDAVLETNTAVEAVAMAYSGDL
                     SDPNEKLYTLDYYLKMAEBIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVH
                     VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAPAHTRRDTGLS
                     LEAVSDLEPYWEAVRGLYLFPESGTPGTPGRVYRHEIPGGQLSNLRAQATALGLADRF
                     ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVI
                     AFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLITEVPEEQAHLDADDSKERRNSLNR
                     LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
                     PDDKGMRNVVANVNGQIRPMRVRDRSVESVTATAEKADSNKGHVAAPFAGVVTVTVA
                     EGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVPAATKVEGDGLIVVVS"

BASE COUNT   788 a   1040 c   1004 g   805 t
ORIGIN

Query Match   97.8%;   Score 3398.8;   DB 1;   Length 3637;
Best Local Similarity   98.6%;   Pred. No. 0;
Matches 3427;   Conservative   0;   Mismatches   47;   Indels   0;   Gaps   0;

```
Qy     1  GTGACTGCTATCACCCTGGCGGTCTTGTTGAAAGATCTTGGTAGCAAACCGGGCGAAATC  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148  GTGACTGCTATCACCCTGGCGGTCTTGTTGAAAGATCTTGGTAGCAAACCGGGCGAAATC  207

Qy    61  CACACATCTCAAGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGGCGAAATC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208  CACACATCTCAAGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGGCGAAATC  267

Qy   121  GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268  GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  327

Qy   181  CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCCGATTGCTACT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328  CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCCGATTGCTACC  387
```

```
Qy  241 GAAGGCTCACCAGTCAAGGCTACCTGGACACTCGATCGATGAAATTATCGGTGCAGCTAAAAAA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  388 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447

Qy  301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  448 GTTAAAGCAGATCCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507

Qy  361 CGCGAGTGCGCGGAAAACGGCATTACTTTATTGCCCAACCCCAGAGGTTCTTGATCTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  508 CGCGAGTGTGCGGAAAACGGCATTACTTTTATTGCCCAACCCCAGAGGTTCTTGATCTC 567

Qy  421 ACCGGTGATAAGTCTCGTGCGTAACCGGCGGAAGAAGGCTGTCTGCCAGTTTTGGCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  568 ACCGGTGATAAGTCTCGCGCGGTAACCGGCGGAAGAAGGCTGGTCTGCCAGTTTTGGCG 627

Qy  481 GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  628 GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687

Qy  541 CCCATCTTGTAAGGCAGTTGCCCGGTGGTGGCCGACCGGGTATGCCGGTATGGCTTCTTCA 600
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  688 CCCATCTTTGTGAAGGCAGTTGCCCGGTGGTGGCCGACCGGGTATGGCCGGTATGGCTTCA 747

Qy  601 CCTGATGAGCTCCGCAAATTGGCAACAGAGACATCTCGTGAAGCTGAAGCGGCATTCGGC 660
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  748 CCTGATGAGCTTTGCTTGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807

Qy  661 GACGGTTCGTATATGTCGGATGCGTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  808 GATGGCGCGGTATATGTCGGATGCGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867

Qy  721 CTTGGCGATCGACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCACGCGT 780
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCACGCGT 927

Qy  781 CGTCACCAAAAAGTTGTCGCAACCGTGACTGAAGAAGTCACCGACATTGGATCCAGAACGTGCGTGAT 840
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  928 CGTCACCAAAAAGTTGTCGCAACCGTGACTGAAGAAGTCACCGACATTGGATCCAGACCACAACGTGCGTGAT 987

Qy  841 CGCATTTGTGCGGATGCAGTGAAGTCTCGCCGCTCGACACGTGGTTACCAGGCGCGGGAACC 900
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  988 CGCATTTGTGCGGATGCAGTGAAGTCTCGCCGCTCGACACGTTGGTTACCAGGGGCGGGAACC 1047

Qy  901 GTGGAATTCTTGGTCGATGAAAAAGGCAACCACCGTTTCATCGAAAATGAACCCACGTATC 960
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  1048 GTGGAATTCTTGGTCGATGAAAAAGGCAACCACCGTTTCATCGAAAATGAACCCACGTATC 1107

Qy  961 CAGGTTGAGCACACCCGTCGACTGAAGAAGTCACCGAGTTGGACCTGGTGAAGGCGCGAGATG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  1108 CAGGTTGAGCACACCCGTCGACTGAAGAAGTCACCGAGTTGGACCTGGTGAAGGCGCAGATG 1167

Qy  1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCCAAGATAAGATCAAGACC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCCAAGATAAGATCAAGACC 1227

Qy  1081 CACGGTGAGCAGCACTGCAGTGCCGCTCACCGGCCATCACCAGGCGAAGATCCAAACAACACGGCTTCCCGCCCA 1140
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1228 CACGGTGAGCAGCACTGCAGTGCCGCCATCACCAGGCGAAGATCCAAACAACACGGCTTCCCGCCCA 1287

Qy  1141 GATACCGGAACTATCACCGGCGGAAATCACCGCTCACCGGTTCGTCTTGACGGT 1200
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1288 GATACCGGAACTATCACCGGCGGAAATCACCGCTCACCGGTTCGTCTTGACGGT 1347

Qy  1201 GCAGCTCAGCTGCGGTGGCGAAATCACCGGCGAAATCTTCAATGCCATGTCGTGCTGTTGCTTCGTGCTGGTCTTGAAAAATGACC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 GCAGCTCAGCTGCGGTGGCGAAATCACCGGCGAAATCTTCAATGCCATGTCGTGCTGTTGCTTCGTGCTGGTCTTGAAAAATGACC 1407

Qy  1261 TGCCGTTGACGGTTCCGACTTTGCACTTGCTCGTGTTGCTGCTGCACCAGCGGCGCGGTTGGCTCAGGTTC 1320
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1408 TGCCGTTGACGGTGTTCCGACTTTGAAACTGCTTCGTTCGTGCTGCACCAGCGGCGCGGTTGGCTCAGTTC 1467

Qy  1321 ACCGGTGCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAGAGGAC 1380
         |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1468 ACCGGTGCTCGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC 1527

Qy  1381 TTCACTTCCAAGCGCATCGGCCATCGCCGGATTTATCGCGATTCATTGCGGATCACCCACCTTCTTCAGGCT 1440
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1528 TTCACTTCCAAGCGCATCGGCCATCGCCGGATTCATTGCGGATCACCCACCTTCTTCAGGCT 1587

Qy  1441 CCACCTGCGGATGATGAGCAGCAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1588 CCACCTGCTGATGATGAGCAGGAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647

Qy  1501 AAGCCTCATGGTGTTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1648 AAGCCTCATGTGTCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707

Qy  1561 AAGGGACTGCCACTGCCACGCGGTTCCGCGACGCCGTGAAGCAGTTGGCCCAGCCGCG 1620
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1708 AAGGGATCTGCCACTGCCACGCGGTTCCGCGACGCCGTGAAGCAGTTGGCCCAGCCGCG 1767

Qy  1621 TTTGCTGCGTCGTGATCTCCGTGAGCAGGAGGACGCACTGGCAGTTACTGCAGTACCACCTTCCGCGAT 1680
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1768 TTTGCTCGTGATCTCCGTGAGCAGGAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1827

Qy  1681 GCACACCAGTCCTTGCTGCCAGTCCGTCATTCGCACTGGACCTGCGGGCAGAG 1740
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1828 GCACACCAGTCTTTGCTTGCCGAGTCCGTCATTCGCACTGGACCTGCGGGCAGAG 1887

Qy  1741 GCCGGTCGCAAAGTCGACTCCTGAGCTTTGTCCGTGAGGCTGGGCGGCGGCGACCTAC 1800
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1888 GCCGGTCGCAAAGTCGACTCCTGAGCTTTTGTCCGTGAGGCCTGGGCGGCGGCGACCTAC 1947

Qy  1801 GATGTGGCGATGCGTTTCTCTTTGAGGATCCGTGGGACACGGCTCGACAGGCTCGCGGAG 1860
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1948 GATGTGGCGATGCGTTTCCTTCTTTGAGGATCCGTGGGACAGGCTCGACAGCTCGCGGAG 2007

Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTGGGATACACC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTGGGATACACC 2067

Qy  1921 CGTACCCAGACTCCGTCGCCGCGTTGTTGGTTAAGGAAGCTGCCAGTCCCGGCGTGGAC 1980
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2068 CGTACCCAGACTCCGTCTGCCGCGTTGTTGGTTAAGGAAGCTGCCAGTCCCGGCGTGGAC 2127

Qy  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGATCGTTCCAGCACGCA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGATCGTCCAGCACGCA 2187

Qy  2041 GTCCTGGAGACCAACACCGGCGGAGGTGGCTATTCGTGTTATTCGTGATCTCTCT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2188 GTCCTGGAGACCAACACCGGCGGAGGTGGCTATGGCTTATTCGTGTTATTCGTGATCTCTCT 2247

Qy  2101 GATCCAAATGAAAAAGCTCTACACCCTGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2248 GATCCAAATGAAAAAGCTCTACACCCTGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307

Qy  2161 AAGTCTGGCGCGTCACATTCTGGCCATTACCACGGAAGATCCATTGGCGTTCGCTTCGCGCTGCG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2308 AAGTCTGGCGCGTCACATTCTGGCCATTACCACTGGAAGATATGGCGTTCGCTCGCGCTGCG 2367

Qy  2221 GTAACCAAGCTGGTCACCGCACTGGCTCCGCACTGTGTTGCTGCTACACCTGCCAGTGCACACCC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2368 GTAACCAAGCTGGTCACCGCACTGGCTCCGCACTGTGTTGCTGCTACACCTGCCAGTGCACACCC 2427

Qy  2281 CACGACACTGCGGGTGGCTTCCGACCACTGTCTGGCGGCACCAGCGGCCATCCCTGTCTGCC 2340
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2428 CACGACACTGCGGGTGGCTTCCGACCACTGTCTGGCGGCACCAGCGGCCATCCCTGTCTGCC 2487

Qy  2341 GCTGTTGACGGTTGCTCCGCACCACTGTTGGCGGCACCACCTCCCGACCGATACCGGTTTGACCGGTTGAGTTC 2400
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  2488 GCTGTTGACGGTTGCTTCCGCACCACTGTTGGCGGCACCACCTCCCGACCGATACCGGTTTGAGCCTGAGTTC 2547

Qy  2401 ATTGGTTGCTGCATTCGGCGCACCCGTCCGGATACAGGGTTTGCTGCGTTGGCGGTTCT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||
```

LOCUS   CGPYC   3728 bp   DNA   linear   BCT 08-MAY-1998
DEFINITION   Corynebacterium glutamicum pyc gene.
ACCESSION   Y09548
VERSION   Y09548.1 GI:2879822
KEYWORDS   pyc gene; pyruvate carboxylase.
SOURCE   Corynebacterium glutamicum.
ORGANISM   Corynebacterium glutamicum
             Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
             Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE   1 (bases 1 to 3728)
AUTHORS   Peters-Wendisch,P.G., Kreutzer,C., Kalinowski,J., Patek,M., Sahm,H.
             and Eikmanns,B.J.
TITLE   Pyruvate carboxylase from Corynebacterium glutamicum:
             characterization, expression and inactivation of the pyc gene
JOURNAL   Microbiology 144 (Pt 4), 915-927 (1998)
MEDLINE   98240228
PUBMED   9579065
REFERENCE   2 (bases 1 to 3728)
AUTHORS   Peters-Wendisch,P.G.
TITLE   Direct Submission
JOURNAL   Submitted (21-NOV-1996) P.G. Peters-Wendisch, Institut fuer
             Biotechnologie 1, Forschungszentrum Juelich GmbH, Juelich, D-52425,
             FRG
FEATURES   Location/Qualifiers
     source   1..3728
                     /organism="Corynebacterium glutamicum"
                     /strain="ATCC 13032"
                     /db_xref="taxon:1718"
     gene   109..3587
                     /gene="pyc"
     mRNA   109..>3587
                     /gene="pyc"
                     /evidence=experimental
     CDS   165..3587
                     /gene="pyc"
                     /EC_number="6.4.1.1"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="CAA70739.1"
                     /db_xref="GI:2879823"
                     /db_xref="SPTREMBL:O54587"
                     /translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQL
                     ARECAENGITFIGPTPEVLDTLGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAAFGDGAVYVERAVINPQH
                     IEVQLIGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDEPLRDRICADAVKFCRSI
                     GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEVTEVDLVKAQMRLAAGATLKEL
                     GLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
                     TGFIADHPHLLQAPPADDEQQRILDYLADVVNKPHGVKPKDVAAPIDKLPNIKDLPL
                     PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVA
                     KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
                     YPDSVCRAFVKEAASSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDL
                     SDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTKLVTALRREFDLPVH
                     VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAHTRRDTGLS
                     LEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSNLRAQATALGLADRF
                     ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPQKYDIPDSVI
                     AFLRGBLGNPPGGWPEPLRTRALEGRSEGKAPLTEVPEEEQAHLDADDSKERRNSLNR
                     LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
                     ACGVTGVTGVTGVGAAGVTGVGAAGVGAGVAGVGAGVGAGVGAGVGVAPPVTVTVA
                     BGDEVKAGDAVAIIRAMKMEATITASVDGKIDRVVPAATKVEGGDLIVVVS"

BASE COUNT   813 a   1060 c   1034 g   821 t
ORIGIN

Query Match   97.8%; Score 3398.8; DB 1; Length 3728;
Best Local Similarity   98.6%; Pred. No. 0;
Matches 3427; Mismatches   0; Conservative   0; Indels 47; Gaps 0;

Qy   1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   114 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 173

RESULT 5
CGPYC

```
Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCGAAGATTGAACCGGTTGTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3474 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCGAAGATTGAACCGGTTGTG 3533

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGTGGCGACTTGATCGTGCGTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3534 GTTCCTGCTGCAACGAAGGTGGAAGTGGCGACTTGATCGTGCGTTCCTAA 3587
```

RESULT 6
LOCUS   A97276   3728 bp   DNA   linear   PAT 26-JAN-2000
DEFINITION Sequence 1 from Patent WO9918228.
ACCESSION   A97276
VERSION   A97276.1 GI:6780661
KEYWORDS   .
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 3728)
  AUTHORS   Peters-Wendisch,P. and Eikmanns,B.
  TITLE   METHOD FOR MICROBIAL PRODUCTION OF AMINO ACIDS OF THE ASPARTATE
            AND/OR GLUTAMATE FAMILY AND AGENTS WHICH CAN BE USED IN SAID METHOD
  JOURNAL   Patent: WO 9918228-A 1 15-APR-1999;
            PETERS WENDISCH PETRA (DE); EIKMANNS BERND (DE)
FEATURES   Location/Qualifiers
     source   1..3728
                     /organism="unidentified"
                     /db_xref="taxon:3264"
BASE COUNT   813 a   1060 c   1034 g   821 t
ORIGIN

Query Match   97.8%; Score 3398.8; DB 6; Length 3728;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3427; Conservative   0; Mismatches 47; Indels   0; Gaps   0;

```
Qy    1 GTGACTGCTATCACCCTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  114 GTGACTGCTATCACCCTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 173

Qy   61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAACCGCGCGAAATC 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  174 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAACCGCGCGAAATC 233

Qy  121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCTATTACCCC 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  234 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCTATTACCCC 293

Qy  181 CGTGAAGATCGGGATCATTCCACCGGTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  294 CGTGAAGATCGGGATCATTCCACCGGTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 353

Qy  241 GAAGGCTCACCAGTGCAAGGCGGTACCTGGACATCGATGAAATATCGGTGCAGCTAAAAAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  354 GAAGGCTCACCAGTGCAAGGCGGTACCTGGACATCGATGAAATATCGGTGCAGCTAAAAAA 413

Qy  301 GTTAAAGCAGATGCTATTTACCGGGATAAGGGTATGCCAACGTC 360
        ||||||||||||||||||||||||||||||||||||||||||||
Db  414 GTTAAAGCAGATGCTATTTACCGGGATAAGGGTATGCCAACGTC 473

Qy  361 CGCGAGTGCGCGGGAAAACGGCATTACTTTTATTGGCCAACCCCAGAGGTTCTTGATCTC 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  474 CGCGAGTGTGCGGGAAAACGGCATTACTTTTATTGGCCAACCCCAGAGGTTCTTGATCTC 533

Qy  421 ACCGGTGATAAGTCTCGCGCGGTAACCGCGCCGGAGAAGGCGGTCTGCCAGTTTTGGCG 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  534 ACCGGTGATAAGTCTCGCGCGGTAACCGCGCCGGAGAAGGCGGTCTGCCAGTTTTGGCG 593

Qy  481 GAATCCACCCCGAGCAAAAACATCGTTAAAAGGCTGAAGGCCAGACTTAC 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  594 GAATCCACCCCGAGCAAAAACATCGTTAAAAGGCTGAAGGCCAGACTTAC 653
```

```
Db  2334 GTAACCAAGCTGGTCACCGACTGCCGCCGTGAATCTGCCAGTGCACGTGCACACC 2393

Qy  2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTCGCAGCTCAAGCTGGTGCAGAT 2340
         ||||||||||||||||||||||  ||||  |||||||||||||||||||||||||||||||
Db  2394 CACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTCGCAGCTCAAGCTGGTGCAGAT 2453

Qy  2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGCACCACTGTCTGCGCACCACTGTCTGCC 2400
         ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2454 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGCACCACTGTCTGCGCACCACTGTCTGCC 2513

Qy  2401 ATTGTTGCTGCTGCATTCGGCGCACCCGTCGCGATACCGGTTTGAGCCTGAGGCTGTTTCT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2514 ATTGTTGCTGCTGCATTCGGCGCACCCGTCGCGATACCGGTTTGAGCCTGAGGCTGTTTCT 2573

Qy  2461 GACCTTGCGAGCCTACTGGGAAGCTGTGGCGCGACTGTACCTGCCATTTGAGTCTGGAACC 2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2574 GACCTTGCGAGCCTACTGGGAAGCTGTGGCGCGACTGTACCTGCCATTTGAGTCTGGAACC 2633

Qy  2521 CCAGGGCCAACCGTGCGTCTACCGGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2634 CCAGGGCCAACCGTGCGTCTACCGGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2693

Qy  2581 CGTGCACAGGCCACCCCACTGGGCCTTGCTGCTTGCTGCTCGATGCTCATCGATGCGAGACAACTAC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2694 CGTGCACAGGCCACCCCACTGGGCCTTGCTGCTTGCTGATCGATGCTCATCGATGCGAGACAACTAC 2753

Qy  2641 GCAGGCGTTAATGAGATGCTGGGACGGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2754 GCAGGCGTTAATGAGATGCTGGGACGGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2813

Qy  2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCCGAGAC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db  2814 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT 2873

Qy  2761 CCACAAAAGTACGACATCCAGACTCGTGTCATCGCGGTTCCTGCGCGAGCTTGGTAAC 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2874 CCACAAAAGTACGACATCCAGACTCGTGTCATCGCGGTTCCTGCGCGAGCTTGGTAAC 2933

Qy  2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGACCCCGCCACTGCGAAGGCCGCTCCGAAGGC 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2934 CCTCCAGGTGGCTGGCCAGAACCACTGCGACCCCGCCACTGCGAAGGCCGCTCCGAAGGC 2993

Qy  2881 AAGGCACCTCTGACCGGAAGTTCCTGAGGAAGCAGGCGCACCTGACGCTGATGATTCC 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2994 AAGGCACCTCTGACCGGAAGTTCCTGAGGAAGCAGGCGCACCTGACGCTGATGATTCC 3053

Qy  2941 AAGGAACGTCGCAACAGACCTCAACCGCCTGCTCGTTCCCGAGAGCCAACGAAGAGTTCCTC 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3054 AAGGAACGTCGCAATGCCTCAACCGCCTGCTCGTTCCCGAGAGCCAACCGAAGAGTTCCTC 3113

Qy  3001 GAGCACCGTGCCGCTGCGGCCCGCCACCTTCGCTGATGATCGTGAATTCTTCTACGGA 3060
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  3114 GAGCACCGTGCCGCTGCGGCCCGCCACCTTCGCTGATGATCGTGAATTCTTCTACGGC 3173

Qy  3061 CTGGTCGAGGGCCGAGACTTTGATCCGGCAGATAAGGGTTGCCGCACCCGACCTGCTTGTT 3120
         |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  3174 CTGGTCGAGGGCCGAGACTTTGATCCGGCAGATAAGGGTTGCCGCACCCGACCTGCTTGTT 3233

Qy  3121 CGCCTTGGATGCGATCTCTGAGCCAGACGCGATAAGGCCATTGCTTGCCTGCCAACGTC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3234 CGCCTTGGATGCGATCTCTGAGCCAGACGCGATAAGGCCATTGCTTGCCTGCCAACGTC 3293

Qy  3181 AACGGCCAGATCCGCGAAAACGGCCATTGCTTGCGTGTCGTGTCACCCGCAACC 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3294 AACGGCCAGATCCGCGAAAACGGCCATTGCTTGCGTGTCGTGTCACCCGCAACC 3353

Qy  3241 GCAGAAAAGGCCAGATTCCTCCAACAGGGCCCAACAGGGCCATGTTGCTGCCACCATTCGCTGGTTGTTGTC 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3354 GCAGAAAAGGCCAGATTCCTCCAACAGGGCCCAACAGGGCCATGTTGCTGCCACCATTCGCTGGTTGTTGTC 3413

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3414 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3473
```

```
Qy  541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCCGGACGCGGTATGCGCTTTGTTTCTTCA  600
        |||||||||| |||||||||||||||||||||||||||||||||||||| ||||| |||||
Db  654 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCCGGACGCGGTATGCGGTTTGTTGCTTCA  713

Qy  601 CCTGATGAGCTTCGCCAAATTGCAACAGAAGCATCTGTGAAGCGGCATTCGGC  660
        ||||||||||||||||||||| ||||||||||||||||||||| |||| |||||
Db  714 CCTGATGAGCTTCGCCAAATTAGCAACAGAAGCATCTGTGAAGCTGAAGCGGCTTTCGGC  773

Qy  661 GACGGTTCGGTATATGTCGACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC  720
        |||  | ||||||||||||||||||||||||| |||||||| |||||||||||||||||
Db  774 GATGGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC  833

Qy  721 CTTGGGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCGAGCGT  780
        |||||||||| ||||||||||||||||||| |||||||||||||||||||||||| ||| |
Db  834 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  893

Qy  781 CGTACCCAAAAAGTTGTCGAAATTGCGCCACAGCAATTGGATCCAGAACTGCGTGAT  840
        ||||||||||||| ||||| ||||| ||||||||||| ||| |||||||||||| |||
Db  894 CGTACCCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAC  953

Qy  841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC  900
        ||||||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  954 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC  1013

Qy  901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC  960
        |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1014 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTCCATCGAAATGAACCCACGTATG  1073

Qy  961 CAGGTTGAGCACACCGTGACTGACTGCAACTGGTGAGAAGTCACCGAGGTGGACCGCGATG 1020
        ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 1074 CAGGTTGAGCACACCGTGACTGACTGCAACTGGTGAGAAGTCACCGAGGTGGACCGCGATG  1133

Qy 1021 CGCTTGGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1134 CGCTTGGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1193

Qy 1081 CACGGTGAGCACTGCGAGTCGCGCCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1194 CACGGTGAGCACTGCGAGTCGCGCCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1253

Qy 1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT  1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1254 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT  1313

Qy 1201 GCAGCTCAGCTCCGGTGGCGAAATCACCGGCACACTTTGACTCTGTGTGAAAATGACC  1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1314 GCAGCTCAGCTCCGGTGGCGAAATCACCGGCACACTTTGACTCTGTGTGAAAATGACC  1373

Qy 1261 TGCCGCTGGTTCCGACTTTGAAACTGCGTGTTGCTCGTGCAGCGGCGCGTTGGCTGAGTTC 1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1374 TGCCGCTGGTTCCGACTTTGAAACTGCGTGTTGCTCGTGCACAGCGGCGCGTTGGCTGAGTTC 1433

Qy 1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGTTGCTGCGGAAGAGGAC  1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1434 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGTTGCTGCGGAAGAGGAC  1493

Qy 1381 TTCACTTCCAAGGCATCGCCACCGGATTTATCGGGATCACCCACACCTTCCTTCAGGCT  1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1494 TTCACTTCCAAGGCATCGCCACCGGATTTATCGGGATCACCCACACCTTCCTTCAGGCT  1553

Qy 1441 CCACCTGCCGATGATGAGCACAGGGACCGGATCCGGATTACTTGGCAGATGTCACCGTGAAC 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1554 CCACCTGCTGATGATGAGCACAGGGACCGCATCCGGATTACTTGGCAGATGTCACCGTGAAC 1613

Qy 1501 AAGCCTCATTGGTGTGCCTCCAAGGGATGTTGCAGCACCAATTGCGATAAAGCTGCCCAACATC 1560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1614 AAGCCTCATTGGTGTGCCTCCAAGGGATGTTGCAGCACCAATCGATAAAGCTGCCTAACATC 1673

Qy 1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCTAGCTG  1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1674 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCG 1733

Qy 1621 TTTGCTCGTCGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 1734 TTTGCTCGTGATCTCCGTGAGCAGGACCGACCCGAGTCCGCACCACTGGCAGTACTGATACCACCTTCCGCGAT 1793

Qy 1681 GCACACCAGTCTTTGCTTGCTTGCGAACCGGACACCGAGTCCGCTCATTCGGACACTACTGAAGCCTCGCCGGCAGAG 1740
Db 1794 GCACACCAGTCTTTGCTTGCTTGCGAACACCGAGTCCGCACTGCACTCGGAGGCTCGGGCCGGCGGCGGACCTAC 1853

Qy 1741 GCCGTCGCAAAGTCGACTCCCTGAGCTTTTGTCCGTGAGGCGTCGGGGCGGGCGCGGACCTAC 1800
Db 1854 GCCGTCGCAAAGCTGACTCCCTGAGCTTTTGTCCGTGAGGCCTCGGGGGCGGGCGGACCTAC 1913

Qy 1801 GATGTGGCGATGCGGTTCCTCTTTGAGGATCGTCGTGGGACAGGCTCGACGAGCTGCGCGAG 1860
Db 1914 GATGTGGCGATGCGGTTTCCTCTTTGAGGATCCGTCGTGGGGACAGGCTCGACGAGCTGCGCGAG 1973

Qy 1861 GCGATCCGGAATGTAAACATTCAGATGCTGCTTCGCGCGCCCGGCAAACACCGTGGGATACACC 1920
Db 1974 GCGATCCGGAATGTAAACATTCAGATGCTGCTTCGCGCGCCCCAACACCGTGGGATACACC 2033

Qy 1921 CGTACCCAGACTCCGTGCGACGCGCTTAACGACGTCTCCCAGAGTCGCCAGCAATCGACGCA 1980
Db 2034 CCGTACCCAGACTCCGTGCGACGCGCTTAACGACGTCTCCCAGATGCGTCCCAGCAATCGACGCA 2093

Qy 1981 ATCTTCCGCATCTTCGACGCGCTTAACCGCGGTTGGCTATATGCGCTTATTCGTGATCTCTCT 2040
Db 2094 ATCTTCCGCATCTTCGACGCGCTTAACCGCGGTTGGCTATATGCGCTTATTCGTGATCTCTCT 2153

Qy 2041 GTCCTGGAGACCCAAGACCGCGGTTGGCTATGGCTTATTCGTGATCTCTCT 2100
Db 2154 GTCCTGGAGACCCAAGACACCGCGGTTGGCTATGGCTTATTCGTGATCTCTCT 2213

Qy 2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGGCAGAGAGATCGTC 2160
Db 2214 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGGCAGAGAGATCGTC 2273

Qy 2161 AAGTCTGGCGCTCACATTCTGGCCATTAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2220
Db 2274 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG 2333

Qy 2221 GTAACCAAGCTGGTCAGTTGGCCAGTGGCACTGCGCGCGGTGAATTCGAATTCGTCAAGCTGGTGGTCGCAGAT 2280
Db 2334 GTAACCAAGCTGGTCAGTTGGCCAGTGGCACTGCGCGCGGTGAATTCGATCTGCGCGAGTCCAGTGCACACC 2393

Qy 2281 CACGACACTGCGGGTGCGGTGGCCAACCTGCGCCACCTCCGTGTCTGCC 2340
Db 2394 CACGACACTGCGGGTGCGGTGGCCAACCTACTTTGTCGGACCACCTCCAGCGCATCCCGTCTGCC 2453

Qy 2341 GCTGTTGACGGTGCTTCCGCACACTGTCCTGCGGCCACCACTCCCAGCCATCCCGTCTGCC 2400
Db 2454 GCTGTTGACGGTGCTTCCGCACACTGTCCTGCGGCCACCACTCCCAGCCATCCCGTCTGCC 2513

Qy 2401 ATTGTTGCTGCATTCGGCGCACACCCGTCGCGGCAGATACCGGTTTGAGCCTGCGTTTCT 2460
Db 2514 ATTGTTGCTGCATTCGGCGCACACCCGTCGCGGCAGATACCGGTTTGAGCCTGTTTCT 2573

Qy 2461 GACCTGCAGACCTACTGGAAGCGTGTGCGCGGACTGTACCTGGAGTGTTGAGTCTGGAACC 2520
Db 2574 GACCTGCAGACCTACTGGAAGCGTGTGCGCGGACTGTACCTGGAGTGTTGAGTCTGGAACC 2633

Qy 2521 CCAGGCCCAACCGGTCGCGTCTACCGGCGACAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
Db 2634 CCAGGCCCAACCGGTCGCGTCTACCGGCGACAAATCCCAGGCGGACAGTTGTCCAACCTG 2693

Qy 2581 CGTGCACAGGCCACCGACACTGGGCCTTGCTGTGATCGTCATCGAAGACAACTAC 2640
Db 2694 CGTGCACAGGCCACCGACACTGGGCCTTGCTGTGATCGTCATCGAAGACAACTAC 2753

Qy 2641 GCAGCCGGTAATGAATGAGATGGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
Db 2754 GCAGCCGGTAATGAATGAGATGGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2813

Qy 2701 GGCGACCTTGCGCACTCCACCTGGTTCGGCGGTCGTAGATCCAGACAGACTTTGCTGCAGAC 2760
```

```
Db   2814  GGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT  2873
Qy   2761  CCACAAAGTACGACATCCCAGACTTCTCGCGCGGCGAGCTTGGTAAC              2820
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   2874  CCACAAAGTACGACATCCCAGACTTCTCGCGCGGCGAGCTTGGTAAC              2933

Qy   2821  CCTCCAGGTGGCTGGCCAGAACCACTGCGACACCCGGCACCCGGCACCCGGAGGCCCGCTCCGAAGGC  2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2934  CCTCCAGGTGGCTGGCCAGAACGACCCACTGCGACACCCGGCACCCGGAGGCCCGCTCCGAAGGC  2993

Qy   2881  AAGGCACCCTCGACGGAAGTTCCTGAGGAAGCAGGCGCACCTGAGCGCTGATGATTCC  2940
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2994  AAGGCACCCTCGACGGAAGTTCCTGAGGAAGCAGGCGCACCTGAGCGCTGATGATTCC  3053

Qy   2941  AAGGAACGTCCCAACAGCCTCAACCCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC  3000
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   3054  AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTC  3113

Qy   3001  GAGCACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGA  3060
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3114  GAGCACCGTCGCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGGC  3173

Qy   3061  CTGGTCGAGGGCCCGGAGACTTGATCCGCCTGCCAGATGTGCGCACCCACTGCTTGTT  3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3174  CTGGTCGAAGGCCCGGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCCACTGCTTGTT  3233

Qy   3121  CGCTCGGATCGGATCTCTGAGCCAGCAGATAAGGGTATGCGCAATGTTGTGGCCAACGTC  3180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3234  CGCTCGGATCGGATCTCTGAGCCAGCAGATAAGGGTATGCGCAATGTTGTGGCCAACGTC  3293

Qy   3181  AACGGCCAGATCCCGCCCAATGCGTGCGTGACCTCCGTTGAGTCTGTCGTCTGCCAACC  3240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3294  AACGGCCAGATCCGCCCAATGCGTGCGTGACCTCCGTTGAGTCTGTCGTCTGCCAACC  3353

Qy   3241  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGCGTGTGTC  3300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3354  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGCGTGTTGTC  3413

Qy   3301  ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCGTGGAGATGCAGTCGCCAATCATCGAG  3360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3414  ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCGTGGAGATGCAGTCGCCAATCATCGAG  3473

Qy   3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGTTGACCGGCAAGATTGAACGCGTTGTG  3420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3474  GCTATGAAGATGGAAGCAACAACAATCACTGCTTCTCGTTGTTGACCGGCAAAATCGATCGCGGTTGTG  3533

Qy   3421  GTTCCTGCTGCAACGAAGGTGGCAACAAGGTGGCGAGTGGGCGACTTGATCGTCGTTCCTAA  3474
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3534  GTTCCTGCTGCAACGAAGGTGGCAACGAAGGTGGCGAGTGGGCGACTTGATCGTCGTTCCTAA  3587
```

COMMENT   Tel:81-44-829-3031, Fax:81-44-813-1651)
               This sequence is conducted by collaboration of Kyowa Hakko Kogyo
               Co. Ltd. And Kitasato University.
FEATURES   Location/Qualifiers
     source   1..332050
                     /organism="Corynebacterium glutamicum ATCC 13032"
                     /strain="ATCC 13032"
                     /db_xref="taxon:196627"
                     /note="ATCC 13032"
     gene   314..2518
                     /gene="Cgl0654"
     CDS   314..2518
                     /gene="Cgl0654"
                     /note="PF02554:Carbon starvation protein CstA"
                     /codon_start=1
                     /transl_table=11
                     /product="Carbon starvation protein, predicted membrane
                     protein"
                     /protein_id="BAB98047.1"
                     /db_xref="GI:21323420"
                     /translation="MKKSPKMSPTARVGLLVFGVIAAVGWGAIAFSRGETINSVWLVL
                     AAVGSYIIAFSFYARLIEYKVVKPKDQRATPAEYVNDGKDYVPTDRRVLFGHHFAAIA
                     GAGPLVGPVMAAQMGYLPGTIWIILGVIFAGAVQDYLVLWVSTRRRGRSLGQMVRDEM
                     GTVGGAAGILATISIMIIIAVLALIVVNALADSPWGVFSITMTIPIALFMGVYLRYL
                     RPGRVTEVSIIGVALLLLAIVAGGWVADTSWGVEWFTWSKTTLALALIGYGIMAAILP
                     VWLLLAPRDYLSTFMKIGVIGLLAVGILFARPEVQMPSVTSFALENGNPVSGSLFPF
                     LFITIACGALSGFHALISSGTTPKLVEKESOMRLGYGGMLMESFVAMMALTAVILD
                     RHLYFSMNAPLALTCGDPATAAEWVNSIGLTGADITPEQLSEAAESVGESTVISRTCG
                     APTLAFGMSEILSLSGFIGGAGMKAFWYHFAIMFEALFILTTVDAGTRVARFMTDTLGN
                     VPGLRRFKDPSWTVGNWISTVFVCALWGAILLMGVTDPLGGINVLFLFGIANQLLAA
                     IALALVLVVVKKGLYKWAWIPAVPLAWDLIVTMTASWQKIFHSDPAIGYWAQNANFR
                     DAKSOGLTEFGAAKSPEAIDAVIRNTMIQGILSILPAVLVLVVVGAIAVCIKSIRAR
                     AAGTPLETTEEPDTESEFFAPTGFLASSRDKEVQAWVDERYPGGAPVSSGGH"
     gene   2522..2725
                     /gene="Cgl0655"
     CDS   2522..2725
                     /gene="Cgl0655"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical protein"
                     /protein_id="BAB98048.1"
                     /db_xref="GI:21323421"
                     /translation="MMALTHALWKIPRAVWWYLTELMGDTAYSKYVVHLKHHPDAPI
                     PTEREYWRAKYADQDANPGARCC"
     gene   complement(3092..3703)
                     /gene="Cgl0656"
     CDS   complement(3092..3703)
                     /gene="Cgl0656"
                     /codon_start=1
                     /transl_table=11
                     /product="Hypothetical protein"
                     /protein_id="BAB98049.1"
                     /db_xref="GI:21323422"
                     /translation="MSNVINLVLMQAHSTHQVNLNLIPSSDATDQILATFSSLLCCSQ
                     NRRNVISWVGVFRSEERVMEIKLTHRSAVSKRSPPFRGITSWQTENRCTPTIGNWPDLC
                     HRTGRANWSTQHGSCLHRCIVDDAVDDHFLSVRSNLNWVGCNASDLVSQVLLVRELFR
                     RTVGAHFVVNHGWNVTPSQRTNVAIHTNSLKSNVGHCKSCVIL"
     gene   3750..4706
                     /gene="Cgl0657"
     CDS   3750..4706
                     /gene="Cgl0657"
                     /codon_start=1
                     /transl_table=11
                     /product="Uncharacterized protein involved in propionate
                     catabolism"
                     /protein_id="BAB98050.1"
                     /db_xref="GI:21323423"
                     /translation="MWNTIYQAIGQALHTTATRQSRKGAISSWKAFAPAFAGKMSIE
                     AVDRAMRGEGAPSPIWEGEDGVIAWLLSGLDHIYTIPLPAEGEAKRAILDYTKEHSA
                     EYQSQAPIDLARSMGEKLAAQGLIDLRDVDSIVLHTSHHTHYVIGTSNDPQKFDDAS
                     RETLDHSIMYIFAVALEDRAWHHERSYAPERAHRRETIELWNKISTVEDPEWTRRYHS
                     VDPAEKAFGARAVITFKDGTVEDELAVADAHPLGARPFAREQYIQKFRTLAEGVVSE
                     KEQDRFLDAAQRTHELEDLSELNIELDADIILAKAPVIPEGLF"

```
gene    4715..5629
        /gene="Cgl0658"
CDS     4715..5629
        /gene="Cgl0658"
        /codon_start=1
        /product="PEP phosphonomutase and related enzymes"
        /protein_id="BAB98051.1"
        /db_xref="GI:21323424"
        /translation="MFSSAVAPTERRKALRALAALAAPEIARMPGAFSPLAARAIQEAGF
        EGVYVSGAVVAADLALPDIGLTLTEVAHRSRQIARVTDLPVLVDADTGFGEPMSAAR
        TVSELEDAGVAGCHLEDQVNPKRCGHLDGKEVVGTDIMVRRIAAAVNERRDEQFVICA
        RTDAAGVEGIDSAIERAKAYADAGADMIFTEALYSPADFEKFRAAVDIPLLAANTEFG
        KTELLPAQLLEDIGYNAVIYPVTLLRIAMGQVEQALGDIANTGICTDWVDRMQHRSRL
        YELLRIRNEYNAFDQQVFTYSADSYKPIF"

gene    5658..6809
        /gene="Cgl0659"
CDS     5658..6809
        /gene="Cgl0659"
        /EC_number="4.1.3.7"
        /note="PF00285:Citrate synthase"
        /codon_start=1
        /product="Citrate synthase"
        /protein_id="BAB98052.1"
        /db_xref="GI:21323425"
        /translation="MSSATTTDVRKGLYGVIADYTAVSKVMPETNSLTYRGYAVEDLV
        ENCSFEEVFYLLWHGELPTAQQLAEFNERGRSYRSLDAGLLISLIHSLPKEAHPMDVMR
        TAVSVMGTKDSEYFTTDSEHIRKVGHTLLAQLPMVLAMDIRRRKGLDIIAPDSSKSVA
        ENLLSMVFGTGPESPASNPADVRDFEKSLILYAEHSFNASTFTARVITSTKSDVYSAI
        TGAIGALKGPLHGGANEFVMHTMLAIDDPNKAAAWINNALDKNKNVVMGFGHRVYKRGD
        SRVPSMEKSFRELAARHDGEKWVAMYENMRDAMDARTGIKPNLDFPAGPAYHLLGFPV
        DFFTPLFVIARVAGWTAHIVEQYENNSLIRPLSEYNGEEQREVAPIEKR"

gene    complement(6846..7040)
        /gene="Cgl0660"
CDS     complement(6846..7040)
        /gene="Cgl0660"
        /codon_start=1
        /trans1_table=11
        /product="Hypothetical protein"
        /protein_id="BAB98053.1"
        /db_xref="GI:21323426"
        /translation="MYFDVSQDVDDELHFSFELYADQAALDAHRCAPHFYVWREAVA
        KHVLPGSQKNMVGTRMLHHG"

gene    complement(7129..8172)
        /gene="Cgl0661"
CDS     complement(7129..8172)
        /gene="Cgl0661"
        /EC_number="1.1.82"
        /note="PF02615:Malate/L-lactate dehydrogenase"
        /codon_start=1
        /trans1_table=11
        /product="Malate/L-lactate dehydrogenases"
        /protein_id="BAB98054.1"
        /db_xref="GI:21323427"
        /translation="MPEVTVNAQQLTVLCTDILLTKTGVPAADAHLVGDSLVQADLWGH
        PSHGVLRLPWYVRRLHSGAMTHAHVEVLNDLGAVILALDGHNGIGQVLADHARKEAVT
        RAMMFGIGAVSVRNSNHFGTAMYTRKAAAQGCVSIITTNASPAMAPWGGREKRIGTN
        PWSIAAPFGETATVVDIANTAVARGKIYHARQTNMPIPETWAITSEGAPTTDPAEAIN
        GVVLPMAGHKGYA1SFMDVLSGVLTGSQHSTKVHGPYDPTPPGAGHLFIALDVAAP
        RDPQDFDDALSDLVGEVKSTPKAQNTEEIFYPGESEDRAHRKNSAHGISLPEKTWMEL
        QELAIENHVVTHR"

gene    complement(8265..8987)
        /gene="Cgl0662"
CDS     complement(8265..8987)
        /gene="Cgl0662"
        /codon_start=1
        /trans1_table=11
        /product="Transcriptional regulators"
        /note="PF00392:Bacterial regulatory proteins, gntR family"
        /protein_id="BAB98055.1"
        /db_xref="GI:21323428"
        /translation="MVQKDAQASPATRKADQVYTQIRREIEDGTLNPGQRMSEVWLVE
```

```
        HTGASRTPVRDALRRLAADELILEPRQAPMVSPLSLRHIKDLFEFRRIVEVAALEEI
        SVGASKSPRIFGEFSTLAADFRELENSADDADFTADFRRLTSKFPDLVAANTHNQFLG
        RSILSLKPHTTRLRIIAHSDHARLRQSVQEHIEMCEAVASGDLRSAGAACRQHLIHVE
        KSIILTALINADSTGSQGIDIRS"

gene    complement(9094..9798)
        /gene="Cgl0663"
CDS     complement(9094..9798)
        /gene="Cgl0663"
        /codon_start=1
        /trans1_table=11
        /product="Hypothetical membrane protein"
        /protein_id="BAB98056.1"
        /db_xref="GI:21323429"
        /translation="MRFSRVLPALLITTAVSIPTASAATLTADTDKELCIASNTDDSA
        VVTPWNSIEDSVREQRLDELDAQDPGIKAAIESYIAQDNAPTAAELQVRLDAIESGE
        GLAMLLPDDPTLADPNAEBESFKTEYTYDEAKDIISGFSSDPASDVLSQUQQAATTGTR
        TAEIRAEVFADRTDDYNESQTALKEDFQNCIDAIDDARPIPLQYLIGGAIALAVIVL
        GIRAWTNSRKQSKHSQ"

gene    complement(9878..12094)
        /gene="Cgl0664"
CDS     complement(9878..12094)
        /gene="Cgl0664"
        /EC_number="1.1.1.42"
        /note="TIGR00178:monomer_idh: isocitrate dehydrogenase,
        NADP-dependent"
        /codon_start=1
        /trans1_table=11
        /product="Monomeric isocitrate dehydrogenase (NADP+)"
        /protein_id="BAB98057.1"
        /db_xref="GI:21323430"
        /translation="MAKIIWIRTDEAPLLATYSLKPVEAFAATAGIEVETRDISLAG
        RILAQFPPERLTEDQKVGNALAELGELAKTPEANIIKLPNISASVPQLKAAIKELQDQG
        YDIPELPDNATTDEEKDILARYNAVKGSAVNPVLREGNSDRRAPIAVKNFVKKFPHRM
        GEWSADSKTNVATMDANDFRHNEKSIILDAADEVQIKHIAADGTETILKDSLKLLEGE
```

```
Query Match     97.8%;  Score 3398.8;  DB 1;  Length 332050;
Best Local Similarity   98.6%;  Pred. No. 0;
Matches 3427;  Conservative   0;  Mismatches 47;  Indels   0;  Gaps   0;

Qy     1 GTGACTGCTATCACCCTTGCGCTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37210 GTGACTGCTATCACCCTTGCGCTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 37269

Qy    61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGCGAAATC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37270 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGCGAAATC 37329

Qy   121 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37330 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 37389

Qy   181 CGTGAAGATGGGGATCATTCCACCGGTCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37390 CGTGAAGATGGGGATCATTCCACCGGTCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 37449

Qy   241 GAAGGCTCACCAGTCAAGGCGTCAAGGCGTAACCGCCGAAGAAGGCGTGGTCGCAGCTAAAAAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37450 GAAGGCTCACCAGTCAAGGCGTCAAGGCGTAACCGCCGAAGAAGGCGTGGTCGCAGCTAAAAAA 37509

Qy   301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCGTGTCTGAAAATGCCCAGCTTGCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37510 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCGTGTCTGAAAATGCCCAGCTTGCC 37569

Qy   361 CGCGAGTGCGCGGGAAAACGGCATTACTTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37570 CGCGAGTGCGCGGGAAAACGGCATTACTTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 37629

Qy   421 ACCGGTGATAAGTCTCTGCGCGTAACCGCCGAAGAAGACGGCGGTCGTGCTCGCCAGTTTGGCG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 37630 ACCGGTGATAAGTCTCTGCGCGTAACCGCCGAAGAAGACGGCGGTCGTGCTCGCCAGTTTGGCG 37689

Qy   481 GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC 540
          |||||||||||||||||||||||||||||||||
```

```
Db  37690 GAATCCACCCGAGCAAAAACATCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTAC 37749

Qy    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGCGGACGCCGGTATGCGCTTTGTTTCTTCA  600
          ||||||||||||||| |||||||||||||||||||||||| |||||| ||||| ||||
Db  37750 CCCATCTTTGAAGGCAGTTGCCGGTGGCGGACGCGGTATGCGCTTTGTTGTTGCTTCA 37809

Qy    601 CCTGATGAGCTCCGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC  660
          ||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db  37810 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 37869

Qy    661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGACATTGAAGTGCAGATC  720
          ||||| | ||||||||||||||||||||| ||||| |||||| |||||||||||||||
Db  37870 GATGGCGCGGTATATGTCGAACGTGCTGTGATTATTGAAGTGCAGATC 37929

Qy    721 CTTGGCGATCCGACCTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780
          |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  37930 CTTGGCGATCCGACCTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 37989

Qy    781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  37990 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 38049

Qy    841 CGCATTTGTCGGATGCAGTAAAGTTCTGCGCCTCCATTGGTTACCAGGGCGCGGGGAACC  900
          ||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db  38050 CGCATTTGTCGCGGATGCAGTAAAGTTCTGCGCCTCCATTGGTTACCAGGGCGCGGGAACC 38109

Qy    901 GTGGAATTCTTGGTCGATGAAAAGGCAACCACGTTTCATCGAAATGAACCCAGTATC  960
          ||||||||| ||||||||||||||||||||||||| |||||||||||||||||||||
Db  38110 GTGGAATTCCTTGGTCGATGAAAAGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 38169

Qy    961 CAGGTTGAGACACCGTCACTGGTGTGCGCATCACCACGGAAGATCAAGATCAAGACC 1020
          ||||||||||||||| |||||||||||| |||||||||||| ||||||| ||||||||
Db  38170 CAGGTTGAGACACCGTCACCGGTGGTGCGGACCTGGTGCTGACCTGGTGCTGACCTGGTGCTGGACCTGACCCAAGATAAGATCAAGACC 38229

Qy   1021 CGCTTGGCTGCTGCTGCAACCTTGAAGGAATTGGGTCTGACCTGACCCAGATAAGATCAAGACC 1080
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  38230 CGCTTGGCTGCTGCTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 38289

Qy   1081 CACGGTGCAGGACTGCAGTCGTGCCGCATCACCACGGAAGATCAAACAACAACGGCTTCCGCCCA 1140
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  38290 CACGGTGCAGGACTGCAGTCGTGCCGCATCACCACGGAAGATCAAACAACAACGGCTTCCGCCCCA 38349

Qy   1141 GATACCGGAACTATCACCGCGTCACCCGGTACCGCTCACCAGGCGCTGGGCGTTCGTCTTCGTGACGGT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38350 GATACCGGAACTATCACCGCGTCACCCGGTGCGGAGCTGGGCGTTCGTCTTCGTGACGGT 38409

Qy   1201 GCGAGCTCAGCTCCGGTGGCGAAATCACCGACGGATTATCGGTGTGCGAAATGACC 1260
          |||||||||||||||||||||||||||||||||||| || |||||||||||||||||
Db  38410 GCGAGCTCAGCTCCGGTGGCGAAATCACCGACACACTTTATCGGTGTGCGAAATGACC 38469

Qy   1261 TGCCGGTGGTTCCGACTTTGAAACTGTCTGTTCGTGCACAGCGCGCCGTTGGCTGAGTTC 1320
          ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  38470 TGCCGGTGGTTCCGACTTTGAAACTGTCTGTTCGTGCACAGCGCGCCGTTGGCTGAGTTC 38529

Qy   1321 ACCGGTGCTGGTGTTGCAACCAACATTGGTTTCTTTGCGTTGCCGTTGCTGCGGGAAGGAGAC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38530 ACCGGTGCTGGTGTTGCAACCAACATTGGTTTCTTTGCGTTGCCGTTGCTGCGGGAAGGAGAC 38589

Qy   1381 TTCACTTCCAAGCGCATCGGCGTGGCGAAATCACCGGATTATCGGCCACCACCTCCTTCAGGCT 1440
          || ||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||
Db  38590 TTCACTTCCAAGCGCATCGGCGTGGCGAAATCACCGGATTATTGCCGATCACCCGCACCTTCTTCAGGCT 38649

Qy   1441 CCACCTGCGGATGATGAGCAGGAGGACGCCATCTCGGATTCGGCAGATGTCACCGTGAAC 1500
          ||||||||||| ||||||||||||||||||||||||||||| |||||||||||||||||
Db  38650 CCACCTGCGGATGATGAGCAGGAGGACGCCATCTCGGATTCGGCAGATGTCACCGTGAAC 38709

Qy   1501 AAGCCTCAAGGGTGTGCCACTGCCACCGGCGTTCCCGTGACCCGCTTGGCCGTGCTCGCGCG 1560
          |||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38710 AAGCCCTGCTACGTGTGATGAGCAGCTGCTTGCGCGATCATGTTGCAGCTCCTTACGTCAACATC 38769

Qy   1561 AAGGATCTGCCACTGCACCGGCCACTGCCGCGCGGTTCCCGTGACCGGCGCTTGGCCCAGCCGCGCG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38770 AAGGATCTGCCACTGCACCGGCCACTGCCGCGCGGTTCCCGTGACCCGCGCTTGGCCCAGCCGCGCG 38829
```

```
Db   1621 TTTGCTCGTGATCTCCGTCCGTGAGCAGGAGCACTGGCCACTGCCAGTTACTGATTACACCTTCCGCGAT 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38830 TTTGCTCGTGATCTCCGTCCGTGAGCAGGAGCACTGGCCACTGCCAGTTACTGATTACACCTTCCGCGAT 38889

Qy   1681 GCACACCAGTCTTGCTTGCGACCCAGTCCGCTCATTCGCACTGAAGCTGCGGGCAGAG 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38890 GCACACCAGTCTTGCTTGCGACCCAGTCCGCTCATTCGCACTGAAGCTGCGGGCAGAG 38949

Qy   1741 GCCGTGCGAAAGCTGACTCTGAGCTTTGTCGTGGAGGCGTGGGGCGGCGCCGACTTAC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38950 GCCGTGCGCAAAGCTGACTCTGAGCTTTGTCGTGGAGGCGTGGGGGGCGGCGCCGACCTAC 39009

Qy   1801 GATGTGGCGGATGCGTTCCTCTTTGAGGATCGTGGGACAGGCTGCGACGAGCTGCGCGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39010 GATGTGGCGGATGCGTTCCTCTTTGAGGATCGTGGGACAGGCTGCGACGAGCTGCGCGAG 39069

Qy   1861 GCGATGCGAATGTAAACATTCAGATGCTGCTTCGCGCGCGCCAACACCGTGGATACACC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39070 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCGCCAACACCGTGGATACACC 39129

Qy   1921 CCGTACCCAGACTCCGTCTGCGCGCGTTGTTAAGGAAGTGCCAGCTCCGGCGTGAC 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39130 CCGTACCCAGACTCCGTCTGCGCGCGTTGTTAAGGAAGTGCCAGCTCCGGCGTGAC 39189

Qy   1981 ATCTTCCGCATCTTCGACGCGTTAACGACGTCTCCCAGATCGTCCAGCAATCGTGACGCA 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39190 ATCTTCCGCATCTTCGACGCGTTAACGACGTCTCCCAGATCGTCCAGCAATCGACGCA 39249

Qy   2041 GTCCTGAGACAACACCGCGGACGGTGGCTATGGCGTATCGTGGTTGGCTTATTCTGGTGATCTCTCT 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39250 GTCCTGAGACAACACCGCGGACGGTGGCTATGGCGTATGGCTTATTCTGGTGATCTCTCT 39309

Qy   2101 GATCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGCGAGAGATCGTC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39310 GATCCAAATGAAAAGTCTCACACCCTGGATTACTACTTGCTGCGAGCTCAAGCTGTGTCGTC 39369

Qy   2161 AAGTCTCGGCGCTCACATTCTGGCCATTATTCTGGCTGTCGCTTCGCCCAGCTGCG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39370 AAGTCTCGGCGCTCACATTCTGGCCATTATTCTGGCTGTCGCTTCGCCCAGCTGCG 39429

Qy   2221 GTAACCAAGCTGGTCACCGACTGGTCACTGCCAGTTGCACTTGCATCTGCCAGTGCACACCC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39430 GTAACCAAGCTGGTCACCGACTGGTCACTGCCAGTTGCACTTGCATCTGCCAGTGCACACCC 39489

Qy   2281 CACGACACTGCGGGTCGGCCAGTTGGCTGGAAGCTGTGTGCGCGGACTGGTCGGTGCTGAGAT 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39490 CACGACACTGCGGGTCGGCCAGTTGGCTGGAAGCTGTGTGCGCGGACTGGTCGGTGCTGCAGAT 39549

Qy   2341 GCTGTTGACGGTGCTTCGGCACACTGTCTGCGCACCACCTCCCAGCCATCCCAGCGTCTGTCTGCC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39550 GCTGTTGACGGTGCTTCGGCACACTGTCTGCGCACCACCTCCCAGCCATCCCAGCGTCTGTCTGCC 39609

Qy   2401 ATTGTTGCTGCATTCGGCAGCCAACACTGGTTTCTTTGCGTTGCCGTTTGAGCTGTTTCT 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39610 ATTGTTGCTGCATTCGGCAGCCAACACTGGTTTCTTTGCGTTTGAGCTGTTTCT 39669

Qy   2461 GACCTTCGAGCCGTACTGGGAAGCTGTGTGCGCGGACAGTGTACCTGCCATTTGAGTCTGGAACC 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39670 GACCTTCGAGCCGTACTGGGAAGCAGTGTGCGCGGACAGTGTACCTGCCATTTGAGTCTGGAACC 39729

Qy   2521 CCAGGCCCAACCGGTCGGCGTCACCGCCACCGGCGGACGCAGGGAATCCAGGCGGACAGTTGTCCAACCTG 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39730 CCAGGCCCAACCGGTCGGCGTCACCGCCACCGGCGGACGCAGGGACAGTTGTCCAACCTG 39789

Qy   2581 CGTGCACAGGCCACCGGCCTTGCTGATGCGTCATCGCGAAGAACAACTAC 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  39790 CGTGCACAGGCCACCGGCCTTGCTGATCATCGAAGACAACTAC 39849

Qy   2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39850 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 39909
```

```
                    /organism="Corynebacterium glutamicum"
                    /db_xref="taxon:1718"
                    /note="Seq 1 to long (3.309.400) split in 11, seq 7061
                    0.600.001 0.949.980 ."

BASE COUNT    79274 a  90638 c  98727 g  81341 t

ORIGIN

Query Match          97.8%;  Score 3398.8;  DB 6;  Length 349980;
Best Local Similarity 98.6%;  Pred. No. 0;
Matches 3427;  Conservative 0;  Mismatches 47;  Indels 0;  Gaps 0;

Qy        1 GTGACTGCTATCACCCTGCGGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   105160 GTGACTGCTATCACCCTTGCGGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 105219

Qy       61 CACACATCTTCAAGGCTTCCAGCATTCAAAAAGATCTGGTAGCAAACCGGCGGCGAAATC 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105220 CACACATCTTCAAGGCTTCCAGCATTCAAAAAGATCTGGTAGCAAACCGGCGGCGAAATC 105279

Qy      121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105280 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 105339

Qy      181 CGTGAAGATCGGGGATCATTCCACCGCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105340 CGTGAAGATCGGGGATCATTCCACCGCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 105399

Qy      241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGATGCAAATTATCGGTGCAGCTAAAAAA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105400 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGATGCAAATTATCGGTGCAGCTAAAAAA 105459

Qy      301 GTTAAAGCAGATGCTATTTACCCGGGATATGCCTCTGTCTGAAAATGCCCAGCTTGCC 360
            |||||||||||||| ||||||||||||||||   |||  ||||||||||||||||||||
Db   105460 GTTAAAGCAGATGCCATTTACCCGGGATATACGGCTTCGTCTGAAAATGCCCAGCTTGCC 105519

Qy      361 CGCGAGTGCGCGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105520 CGCGAGTGCGCGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 105579

Qy      421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAGAGGCTGGTCTGCCAGTTTTGGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105580 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGAAGAGAGGCTGGTCTGCCAGTTTTGGCG 105639

Qy      481 GAATCCACCCCGAGCAAAAACATCGATGCACATCGTTAAAAGCCGTCGAAGGCCAGACTTAC 540
            |||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   105640 GAATCCACCCCGAGCAAAAACATCGATGCACATCGTTAAAAGGCTGAAGGCCAGACTTAC 105699

Qy      541 CCCATCTTGTAAAGGCAGTTGCCGGTGGCGGTGGCGGACGGCGGTATGCGGCTTTGTTCTTCA 600
            ||||||||||  |||||||||||||||||||||||||||  ||||||||||||| || ||||
Db   105700 CCCATCTTGTGAAGGCAGTTGCCGGTGGCGGTGGCGGACCGCGGTATGCGGCTTTTGTTGCTTCA 105759

Qy      601 CCTGATGAGCTTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCATCGTGAAGCGCTTTCGGC 660
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   105760 CCTGATGAGCTTCCGCAAATTAGCAACAGAAGCATCTCGTGAAGCATCGTGAAGCGCTTTCGGC 105819

Qy      661 GACGGTTCGGTATATGTCGAACGTGTCTGATTAACCCCAGCACATTGAAGTGCAGATC 720
            || ||  ||||||||||||||||||| ||||||||||| ||| |||||||||||||||
Db   105820 GATGGGCCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 105879

Qy      721 CTTGGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
            |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   105880 CTTGGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 105939

Qy      781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAATCGGTGGTGAT 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   105940 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAATCGGTGGTGAT 105999

Qy      841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTTCCATTGGTTACCAGGGCCGCGGGAACC 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106000 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTTCCATTGGTTACCAGGGCCGCGGGAACC 106059

Qy      901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCCGAAATGAACCCACGTATC 960
            ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
```

```
Qy     2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGTGGGTGTAGATCCAGCAGACTTTGCTGCGAGAC 2760
            |||||||||||||||||||||| |||||||||||||||  |||||||||||||||||||||||
Db    39910 GGCGACCTCGCACTCCACCTCGTTGGTGCGGTGGGTGTGGATCCAGCAGACTTTGCTGCGAGAC 39969

Qy     2761 CACAAAAGTACGACATCCCAGACTCTGTCATCGCGGTTCCTCGCGGCCGAGCTTGGTAAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39970 CACAAAAGTACGACATCCCAGACTCTGTCATCGCGGTTCCTCGCGGCCGAGCTTGGTAAC 40029

Qy     2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCGCCCGCGCCACTGAAAGGCCGCTCCGAAGGC 2880
            ||||||||||||||||||||| ||||||||||||  |||||||||||||||||||||||||
Db    40030 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCGCCGCCGCCACTGAAAGGCCGCTCCGAAGGC 40089

Qy     2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGACGAGGCCCACCTCGACGCTGATGATTCC 2940
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40090 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGACGAGGCCCACCTCGACGCTGATGATTCC 40149

Qy     2941 AAGGAACGTCGCAACAGCCCTCAACCGCCCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    40150 AAGGAACGTCGCAATAGCCCTCAACCGCCCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 40209

Qy     3001 GAGCACCGTCGCCGCTTCGGCAACCACCCTGCGCTGGATGATCGTGAATTCTTCTACGGA 3060
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    40210 GAGCACCGTCGCCGCTTCGGCAACCACCCTGCGCTGGATGATCGTGAATTCTTCTACGGC 40269

Qy     3061 CTGGTCGAGGGCCGGCCGAGACTTTGATCCGGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
            |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40270 CTGGTCGAAGGCCGGCCGAGACTTTGATCCGGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 40329

Qy     3121 CGCCTGGATGCGATCTCGAGCCAGACGATAAGGGTATGCCGCAATGTTGTGGCCAACGTC 3180
            |||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||
Db    40330 CGCCTGGATGCCGATCTCTGAGCCAGACGATAAGGGTATGCCGCAATGTTGTGGCCAACGTC 40389

Qy     3181 AACGGCCAGATCCGCCCAATGCGTGTCGTGACCGCTCCGTGAGTCTGTCACCGCAACC 3240
            ||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    40390 AACGGGCCAGATCCGCCCAATGCGTGTCGTGACCGCTCCGTGAGTGTGTCACCGCAACC 40449

Qy     3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCCATTCGCTGGTGTTGTC 3300
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||
Db    40450 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTC 40509

Qy     3301 ACTGTGACTGTTGCTGGAAGGTGATGAGGTCAAGGCTGGAGATGCGCAATCATCGAG 3360
            || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40510 ACCGTGACTGTTGCTGGAAGGTGATGAGGTCAAGGCTGGAGATGCGCAATCATCGAG 40569

Qy     3361 GCTATGAAGATGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420
            |||||||||||||||||||| ||||||||||||||||||||||| || |||||||||||
Db    40570 GCTATGAAGATGAAGCAACAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 40629

Qy     3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGAATCGTCGTTTCCTAA 3474
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40630 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGAATCGTCGTTTCCTAA 40683
```

```
RESULT 8
AX127145
LOCUS       AX127145              349980 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 7061 from Patent EP1108790.
ACCESSION   AX127145 AX114121
VERSION     AX127145.1  GI:14041133
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
ORGANISM    Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 349980)
AUTHORS     Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE       Novel polynucleotides
JOURNAL     Patent: EP 1108790-A 7061 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1..349980
```

```
Db 106060 GTGGAATTCTTGGTCGTGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 106119

Qy   961 CAGGTTGAGCACACCGACTGAAGAAGTCACCGAGGTGGACCTGAAGGCGCAGATG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106120 CAGGTTGAGCACACCGACTGAAGAAGTCACCGAGGTGGACCTGAAGGCGCAGATG 106179

Qy  1021 CGCTTGGCTGCTGGTGCGTGTGCAACCTTGACCGTCTGACCCAAGATAAGATCAAGACC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106180 CGCTTGGCTGCTGGTGCGTGTGCAACCTTGACCGTCTGACCCAAGATAAGATCAAGACC 106239

Qy  1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGCGGAAGATCAAACAACGGCTTCCGCCCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106240 CACGGTGCAGCACTGCAGTGCCGCATCACCACGCGGAAGATCAAACAACGGCTTCCGCCCA 106299

Qy  1141 GATACCGGAACTATCACCGGCTGGCTACCGGCGAGCTGGCGTTCGTTGACGGT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106300 GATACCGGAACTATCACCGGCTGGCTACCGGCGAGCTGGCGTTCGTTGACGGT 106359

Qy  1201 GCAGCTCAGCTCGGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106360 GCAGCTCAGCTCGGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 106419

Qy  1261 TGCCGGTGGTTCCGACTTGAAACTGCTCGTGTTCTTGCGTGCACAGCGGCGGCGTTGGCTGAGTTC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106420 TGCCGGTGGTTCCGACTTGAAACTGCTCGTGTTCTTGCGTGCACAGCGGCGGCGTTGGCTGAGTTC 106479

Qy  1321 ACCGTGTCTGTGTTGTTGCAACCAACATTGGTTCTTGCGGATCGGCACCCACACCTTCAGGCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106480 ACCGTGTCTGTGTTGTTGCAACCAACATTGGTTCTTGCGGATCGGCACCCACACCTTCAGGCT 106539

Qy  1381 TTCACTTCCAAGCGATCGCCACCGGATTTATCGCGATCACCCGGACTGGCAGATGTTCACCGGCT 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106540 TTCACTTCCAAGCGATCGCCACCGGATTTATCGCGATCACCCGGACTGGCAGATGTTCACCGGCT 106599

Qy  1441 CCACCTGGCGGACGCAGGACGCAGGACGCAGCATCCGGATCCGTGAAC 1500
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 106600 CCACCTGGCGGACGCAGGACGCAGGACGCAGCATCCGGATCCGTGAAC 106659

Qy  1501 AAGCCTCATGTGTGCGTCCAAAGATGTTCAGCACCAATCGATAAGCTGCCCAACATC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106660 AAGCCTCATGTGTGCGTCCAAAGATGTTCAGCACCAATCGATAAGCTGCCTAACATC 106719

Qy  1561 AAGGATCTGCCACTGCCACCGCGGTTCCCGTGACCGCCTTGGCCCGAGCCG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106720 AAGGATCTGCCACTGCCACCGCGGTTCCCGTGACCGCCTTGGCCCGAGCCG 106779

Qy  1621 TTTGCTCGTCGTTCCTTTGAGGATCCGTCAGTTACTACTTCACCGTGCCGAT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106780 TTTGCTCGTCGTTCCTTTGAGGATCCGTCAGTTACTACTTCACCGTGCCGAT 106839

Qy  1681 GCAACCAGTCTTTGCTTGCGACCTCGCTCATTCGAAGCCTGAAGCTGCGGAGAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106840 GCAACCAGTCTTTGCTTGCGACCTCGCTCATTCGAAGCCTGAAGCTGCGGAGAG 106899

Qy  1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGGCGCGGCCGACCTAC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106900 GCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGGCGCGGCCGACCTAC 106959

Qy  1801 GATGTGGCGATGGCGTTTCCTTTTGAGGATCCGCAGTGGGAACAGGCTCGACGAGCTGCCGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106960 GATGTGGCGATGGCGTTTCCTTTTGAGGATCCGCAGTGGGAACAGGCTCGACGAGCTGCCGAG 107019

Qy  1861 GCGATGCCGAATGTAAACATTCAGATACGTTGTTGCAACAGCTGCCAACACCGTGGGATACACC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107020 GCGATGCCGAATGTAAACATTCAGATGCTTGTTGCAACAGCTGCCAACACCGTGGGATACACC 107079

Qy  1921 CCGTACCCAGACTCCGTCGCCGCCGCGCGTTTGTTAAGGAAGCTGCCACTCGTCCCGGCGGGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107080 CCGTACCCAGACTCCGTCGCCGCCGCGCGTTTGTTAAGGAAGCTGCCACTCGTCCCGGCGGGAC 107139

Qy  1981 ATCTTCCGCATCTTCGACGCGCGTTAACGACGCTCTCCCAGATGCGTCCAGCAATCGACGCGCA 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107140 ATCTTCCGCATCTTCGACGCGCGTTAACGACGCTCTCCCAGATGCGTCCAGCAATCGACGCGCA 107199

Qy  2041 GTCCTGGAGAGCAACCAACACGTCTTCATCGAGGTGGCTATGGCCGAGTTATTCTGCGTGATCTCTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107200 GTCCTGGAGAGCAACCAACACGTCTTCATCGAGGTGGCTATGGCCGAGTTATTCTGCGTGATCTCTCT 107259

Qy  2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107260 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 107319

Qy  2161 AAGTCTTGGCGCTCACATTCTGCCATTAAGGATATGCGTCGTGCTTCGCCCAGCTGCGG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107320 AAGTCTTGGCGCTCACATTCTGCCATTAAGGATATGCGTCGTGCTTCGCCCAGCTGCGG 107379

Qy  2221 GTAACCAAGCTGGTCACCGACTGCCGTGAATTCGATCGATCTGCCAGTGCACGTGCACACC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107380 GTAACCAAGCTGGTCACCGACTGCCGTGAATTCGATCGATCTGCCAGTGCACGTGCACACC 107439

Qy  2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTCGCACCACCTCCGTGCTCTGCC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107440 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTCGCACCACCTCCGTGCTCTGCC 107499

Qy  2341 GCTGTTGACGGTGCTTCCGCGACCACCTGGAAGCTGTGCGGCGACTGTGTCCATTTGAGTCTGGAACC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107500 GCTGTTGACGGTGCTTCCGCGACCACCTGGAAGCTGTGCGGCGACTGTGTCCATTTGAGTCTGGAACC 107559

Qy  2401 ATTGTTGCTGCATTCGGCGACGTGTGCGGACGTGTTACCTGCCATTTGAGCGCCGAGGCTGTTTCT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107560 ATTGTTGCTGCATTCGGCGACGTGTGCGGACGTGTTACCTGCCATTTGAGCGCCGAGGCTGTTTCT 107619

Qy  2461 GACCTGAGCGCCTTGGAAGCTGTGGGAAGCAGTGTACCGGCGACGAAATCCCAGGCGGAGAACC 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107620 GACCTGAGCGCCTTGGAAGCTGTGGGAAGCAGTGTACCGGCGACGAAATCCCAGGCGGAGAACC 107679

Qy  2521 CCAGGCCCAACCGGTGCGTGCTGCGTACTGGCAGGACGCCCAACTGTTGTCCAACCTG 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107680 CCAGGCCCAACCGGTGCGTGCTGCGTACTGGCAGGACGCCCAACTGTTGTCCAACCTG 107739

Qy  2581 CGTGCACAGGCCACCCGACTGGCCTTGCGTGGACGCCTCATCGAGACTCATCGAAGACAACTAC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107740 CGTGCACAGGCCACCCGACTGGCCTTGCGTGGACGCCTCATCGAGACTCATCGAAGACAACTAC 107799

Qy  2641 GCAGCCGTTAATGAGATGTCCGGTGAACGCAGCTTGGCCGTCACCCAAGGTGTT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107800 GCAGCCGTTAATGAGATGTCCGGTGAACGCAGCTTGGCCGTCACCCAAGGTGTT 107859

Qy  2701 GGCGACCTGCACTGCACCTGGTTGCGGGTGTGCTGGGTGTGCGACGCTTTGCTGCGAC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107860 GGCGACCTGCACTGCACCTGGTTGCGGGTGTGCTGGGTGTGCGACGCTTTGCTGCGAT 107919

Qy  2761 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCTCGCGGCGAGCTTGGTAAC 2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107920 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCTCGCGGCGAGCTTGGTAAC 107979

Qy  2821 CCTCCAGGTGGCTGGCTGGCCAGTGGCCTGCGCACCACTGCGCACGGCCGCTCCGAAGGC 2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107980 CCTCCAGGTGGCTGGCTGGCCAGTGGCCTGCGCACCACTGCGCACGGCCGCTCCGAAGGC 108039

Qy  2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGCAGGCGCACCTGGATGATTCC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108040 AAGGCACCTCTGACGGAAGTTCCTGAGGAGCAGGCGCACCTGATGATTCC 108099

Qy  2941 AAGGAACGTCGCAACAGCCTCAACGCCTGTGTTCCCGAGACTCTGTTCATCGCCGCTGTGTTCCTC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108100 AAGGAACGTCGCAATAGCCTCAACGCCTGTGTTCCCGAGACTCTGTTCATCGCCGCTGTGTTCCTC 108159

Qy  3001 GAGCACCGTCGCCGCGCTTCGGCAACAGCTGCGATCGTGGGATCGTGAATTCTTCTACGGA 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108160 GAGCACCGTCGCCGCGCTTCGGCAACAGCTGCGATCGTGGGATCGTGAATTCTTCTACGGC 108219

Qy  3061 CTGGTCGAGGGCGGAGACTTTGATCCGCGTCCCAGAGTGCGCCACCCCACTGCTTGTT 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108220 CTGGTCGAGGGCGGAGACTTTGATCCGCGTCCCCAGAGTGCGCCACCCCACTGCTTGTT 108279
```

```
KLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDELREAMPNVNIQMLLRGRNTVGYTP
YPDSVCRAFVKEAATSGVDIFRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDL
SDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRPAAATKLVTALRREFDLPVH
VHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSALVAAFAHTRRDTGLS
LEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQATALGLGADRF
ELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAADPOKYDIPDSVI
AFLRGELGNPGGWPEPLRTRALEGRSEGKAPLTEVPEEQAHLDADDSKERRNSLNR
LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISE
PDDKGMRNVVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAAPFAGVVTVTVA
EGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLIVVIS"
```

```
BASE COUNT      801 a    1041 c    1006 g     809 t
ORIGIN
```

```
Query Match        97.7%;   Score 3394;  DB 1;  Length 3657;
Best Local Similarity  98.6%;  Pred. No. 0;
Matches 3424;  Conservative  0;  Mismatches  0;  Indels  50;  Gaps  0;
```

```
Qy     1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCAGTGTCGACT   60
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   100 GTGACTGCTATCACCCTTGGTGGTCTCTTGTTGAAAGGAACAATTACTCAGTGTCGACT  159

Qy    61 CACACATCTTCAAGGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGCGCAAAATC  120
         |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGCGCAAAATC  219

Qy   121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  279

Qy   181 CGTGAAGATCGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT  240
         |||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db   280 CGTGAAGATCGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACT  339

Qy   241 GAAGGCTCACCAGTCAAGGCTACCTGCACATCGATGAAATTATCGGTGCAGCTAAAAAA  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340 GAAGGCTCACCAGTCAAGGCTACCTGCACATCGATGAAATTATCGGTGCAGCTAAAAAA  399

Qy   301 GTTAAAGCAGATGCTATTTACCCGGGAAATGCCCGGCGCGTAACCGCCGAAGAAATGCCCAGCTTGCC  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   400 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTAGTTGGCCAACCGCCGAAGAAATGCCCAGCTTGCC  459

Qy   361 CGGAGTGCGCGGAAAACGCCATTACTTTATTGGCCAACCCCAGAGGTTCTTGATCTC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   460 CGGAGTGCGCGGAAAACGCCATTACTTTATTGGCCAACCCCAGAGGTTCTTGATCTC  519

Qy   421 ACCGGTGATAAGTCTCGTGCGTAACCGCCGGAAGAGGCTGGTCTGCCAGTTTTGGCG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   520 ACCGGTGATAAGTCTCGTGCGTAACCGCCGGAAGAGGCTGGTCTGCCCAGTTTTGGCG  579

Qy   481 GAATCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGGCTGAAGGCCAGACTTAC  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   580 GAATCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGGCTGAAGGCCAGACTTAC  639

Qy   541 CCCATCTTTGTAAAGGCAGTTGCCGGTGCGGACGCGGTATGCGCTTTGTTCTTCA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   640 CCCATCTTTGTAAAGGCAGTTGCCGGTGCGGACGCGGTATGCGCTTTGTTCTTCA  699

Qy   601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCCGGCATTCGGC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   700 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCCGGCATTCGGC  759

Qy   661 GACGGTTCGGTATATGTCGAACGTGCTGGTTGATTAACCCCAGCACATTGAAGTGCAGATC  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   760 GACGGTTCGGTATATGTCGAACGTGCTGGTTGATTAACCCCAGCACATTGAAGTGCAGATC  819

Qy   721 CTTGGGGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGGCTCACTGCAGCGT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   820 CTTGGGGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGGCTCACTGCAGCGT  879

Qy   781 CGTCACCAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCCTGAT  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   880 CGTCACCAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCCTGAT  939
```

```
Qy  3121 CGCCTGCGATGGCGATCCTCGAGCCAGACGAGATAAGGGTATGCGCAATGTTGTCGGCCAACGTC  3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108280 CGCCTGCGATGGCGATCCTCGAGCCAGACGAGATAAGGGTATGCGCAATGTTGTCGGCCAACGTC 108339

Qy  3181 AACGGCCAGATCCGCCCAATGCGTGCGTGCGTGACCGCTCCGTGAGTCGTCGTCACCGCAAAC  3240
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 108340 AACGGCCAGATCCGCCCAATGCGTGCGTGCGTGACCGCTCCGTGAGTCGTCGTCACCGCAAAC 108399

Qy  3241 GCAGAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCCACCATTCGCTGGTGTTGTC  3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108400 GCAGAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCCACCATTCGCTGGTGTTGTC 108459

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG  3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108460 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 108519

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG  3420
         |||||||||||||||||||||||||||||||||||||||| || |||||||||||||||
Db 108520 GCTATGAAGATGGAAGCAACAACAATCACTGCTTCTGTTGACGGCAAAATCGATCGCGTTGTG 108579

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGAGGTGACTTGATCGTCGTTTCCTAA  3474
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108580 GTTCCTGCTGCAACGAAGGTGGAAGGTGGAGGTGACTTGATCGTCGTTTCCTAA 108633
```

```
RESULT 9
AF503915
LOCUS       AF503915           3657 bp    DNA     linear   BCT 13-MAY-2002
DEFINITION  Corynebacterium crenatum strain CD945 pyruvate carboxylase (pyc)
            gene, complete cds.
ACCESSION   AF503915
VERSION     AF503915.1  GI:20531756
KEYWORDS    .
SOURCE      Corynebacterium crenatum.
ORGANISM    Corynebacterium crenatum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 3657)
AUTHORS     Wang,J., Ding,J. and Liu,Y.
TITLE       Cloning and Expression of Pyruvate Carboxylase Gene in
            Corynebacterium crenatum CD945
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 3657)
AUTHORS     Wang,J., Ding,J. and Liu,Y.
TITLE       Direct Submission
JOURNAL     Submitted (18-APR-2002) Center of Microbial Biotechnology,
            Institute of Microbiology, Chinese Academy of Sciences, Beijing
            100080, China
FEATURES             Location/Qualifiers
     source          1..3657
                     /organism="Corynebacterium crenatum"
                     /strain="CD945"
                     /db_xref="taxon:168810"
     gene            151..3573
                     /gene="pyc"
     CDS             151..3573
                     /gene="pyc"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="AAM27458.1"
                     /db_xref="GI:20531757"
                     /translation="MSTHTSSTLPAFKKILVANRGEIAVRAFRAALETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSSPVKAYLDIDEIIGAAKKVKADAIYPGYGFLSENAQL
                     ARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDDIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVSSPDELRKLATEASREAEAAFGDGSVYVERAVINPQH
                     IEVQILGDRTGEVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFCRSI
                     GYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEBVTEVDLVKAQMRLAAGATLKEI
                     GLTQDKIKTHGAALQCRITTEDPNNGFRPDTGIITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIA
                     TGFIGDHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIKDLPL
                     PRGSRDRLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAARAVA
```

```
Qy   841  CGCATTTGTGCGGATGCAGTAAAGTTCTGCGCCTTCCATTGGTTACCAGGGGCGGGAACC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   940  CGCATTTGTGCGGATGCAGTAAAGTTCTGCGCCTCCCATTGGTTACCAGGGGCGGGAACC   999

Qy   901  GTGGAAATTCTTGGTCGGATGAAAAGGGGCAACCACGTTTCATCGAAATGAACCCACGTATC   960
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  1000  GTGGAATTCTTGGTCGGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC  1059

Qy   961  CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGCTGAAGGCGCAGATG      1020
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  1060  CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGCTGAAGGCGCAGATG      1119

Qy  1021  CGCTTGGCTGCTGGTGCAAACCTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC   1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1120  CGCTTGGCTGCTGGTGCAAACCTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC   1179

Qy  1081  CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACAACGGCTTCCGCCCCA   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1180  CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACAACGGCTTCCGCCCCA   1239

Qy  1141  GATACCGGAACTATCACCGGTACCGGTCACCGAGCTGGCGTTCGTCTTGACGGT   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1240  GATACCGGAACTATCACCGGTACCGGTCACCGAGCTGGCGTTCGTCTTGACGGT   1299

Qy  1201  GCAGCTTCAGCTTCGGTGCGGTGCGAAATCACCGCACACTTTGACTCCATGCTCGTGTGAAAATGACC   1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1300  GCAGCTTCAGCTTCGGTGCGGTGCGAAATCACCGCACACTTTGACTCCATGCTCGTGTGAAAATGACC   1359

Qy  1261  TGCCGTTGGTTCCGACTTTGAAAACTGCTGTCGTCGTGCACAGCCGCGCGGCGTTGGCTGAGTTC   1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1360  TGCCGTTGGTTCCGATTTTGAAAACTGCTGTCGTCGTGCACAGCCGCGCGGCGTTGGCTGAGTTC   1419

Qy  1321  ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTGCGGTGCTTGCTGCGGGGAAGAGGAC   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420  ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTGCGGTGCTTGCTGCGGGTGAAGAGGAC   1479

Qy  1381  TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGGCGATCACCCCACCTCCTTCAGGCT   1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480  TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGGCGATCACCCCACCTCCTTCAGGCT   1539

Qy  1441  CCACCTGCGGATGAGCACGAGGACGCACTGGACGCCATCCTGGATTACTTGGCAGATGCAGT   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540  CCACCTGCGGATGAGCACGAGGACGCACTGGACGCCATCCTGGATTACTTGGCAGATGTGAAC   1599

Qy  1501  AAGCCTCATGGTGTGCGCTCCAAAGATGTTGCAGCACCAACATCGATAAAGCTGCCCAACATC   1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1600  AAGCCTCATGGTGTGCGCTCCAAAGATGTTGCAGCACCAATCGATAAAGCTGCCCAACATC   1659

Qy  1561  AAGGATCTGCCACTGCACCGCGGTTTCGGGTGACCCGGCTTGGCCAGCAGCGCG      1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1660  AAGGATCTGCCACTGCACCGCGGTTTCGGGTGACCCGGCTTGGCCAGCAGCACGCG      1719

Qy  1621  TTTGCCGTGATCTCCGTGAGCGAGGACGCACTGGACGCCATGGACGCACTGGACGCCATGGCCGAT   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1720  TTTGCCGTGATCTCCGTGAGCGAGGACGCACTGGACGCCATGGACGCACTGGACGCCATGGCCGAT   1779

Qy  1681  GCACACCAGTCGTTCCTTCTTTGCTGGTGGACTCCGTCATTCGCACCACCAACATCGACACTCGACGCCTGCGGCAGAG   1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1780  GCACACCAGTCGTTCCTTCTTTGCTGGTGGACTCCGTCATTCGCACCACCAACATCGACACTCGACGCCTGCGGCAGAG   1839

Qy  1741  CCTCCAGGTGGCTGGCAGACTCCTGAGCTTTTGTCCGTGGAGGCGTTGGCCGCGCTGGGCCGCGCGACCTAC   1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1840  CCTCCAGGTGGCTGGCAGACTCCTGAGCTTTTGTCCGTGGAGGCGTTGGCCGCGCGGGCCGCGCGACCTAC   1899

Qy  1801  GATGTGGCGGATGGCTTCCTTCTTTTGAGGATCGGACCCGAGTCCGTCATTCGCACTCGAAGCCTGCGGCAGAG   1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1900  GATGTGGCGGATGGCTTCCTTCTTTTGAGGATCGGACCCGAGTCCGTCATTCGCACTCGAAGCCTGCGGCAGAG   1959

Qy  1861  GCGATGCCGAATGTAAACATTCAGATGCGCTGTCGTCTTGCGGGTCCCGGCCGCCAACACCGTGGGATACACC   1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1960  GCGATGCCGAATGTCAACATTCAGATGCGCTGCTTGCGGGTCCCGGCCGCCAACACCGTGGGATACACC   2019

Qy  1921  CCGTACCCGAGACTCCGTCTGCGCGCGTTTGTTAAGGAAGCTGCCACCTCCGCCGGTGGAC   1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2020  CCATACCCGAGACTCCGTCTGCGCGCGTTGTTGTTAAGGAAGCTGCCACCTCCGCCGGTGGAC   2079

Qy  1981  ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCTCCCAGCAATCGACGCGA   2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2080  ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCTCCCAGCAATCGACGCGA   2139

Qy  2041  GTCCTGGAGACCAACACCGCGCCGGTAGCGCCGGTCGCTGAAGGTGGCTATGGCTTATTCTGGTGATCTCCT   2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2140  GTCCTGGAGACCAACACCGCGCCGGTAGCGCCGGTCGCTGAAGTGGCTATGGCTTATTCTGGTGATCTTTCC   2199

Qy  2101  GATCCAAATGAAAAGTCTCACACCCTGATTACTACCTAAAGATGGCAGAGGAGATCGTC   2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2200  GATCCGAATGAAAAGTCTCACACCCTGATTACTACCTGAAGATGGCAGAGGAGATCGTC   2259

Qy  2161  AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCG   2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2260  AAGTCTGGCGCTCACATTCTGGCTATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCA   2319

Qy  2221  GTAACCAAGCTGGTCACCGGACTGGCGTTGAATTCATCTGGTCACTGCCAGTGCACTGCACACC   2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2320  GCCACCAAGCTGGTCACCGGACTGGCGTTGAATTTGATCTGCCAGTGCACTGCACACC   2379

Qy  2281  CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCGAGCTCAAGCTCGGTGCGAT   2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2380  CACGACACTGCGGGTGGCCATTACTGGGAAGCAGTGCGAGCTGGCCATTACCTACTTTGCTGCGAGCTCGGTGCGAT   2439

Qy  2341  GCTGTTGACGGTGCTTCCGCAAGCTGTGCTTCGCGCACCTCTCTGCGGATATACCGGTTGAGCCTTGAGCGCTGTTTCT   2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2440  GCTGTTGACGGTGCTTCGCGCACCTCTCTGCGCCACCTCTCTGCGGATACCGGTTGAGCCTTGAGCGCTGTTTCT   2499

Qy  2401  ATTGTTGCTGCATTGCATTCTGGTTCTTGCGGTGCGGACTGTGCGCCGACTGTGTGGAACC   2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2500  ATTGTTGCTGCATTGCATTCTGGTTTTGCGGTGCGGACTGTGCGCCGACTGTGTCAACCTG   2559

Qy  2461  GACCTCGAGCCGTACTGGGAAGCTGTGGACGGCCACCCAACCAAGTCACCCCATCCTCCAAGGTTGTT   2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2560  GACCTCGAGCCCATACTGGGAAGCAGTGCCGACTGTACCTGCCATCCTCCAAGGTTGTT   2619

Qy  2521  CCAGGCCCAACCGGTGCGTCTACCGGCGCACGAAATCCCAGGCGGACAGTTGTCCAACCTG   2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2620  CCAGGCCCAACCGGTGCGTCTACCGGCGCACGAAATCCCAGGCGGACAGTTGTCCAACCTG   2679

Qy  2581  CGTGCACAGGCCACCGGCCTTGCTGCTTCGAGCTCATCGATCGAGAACGACAACTAC   2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2680  CGTGCACAGGCCACCGGCCTTGCTGCTTCGAACTCATCGAACTCATCGAGAACGACAACTAC   2739

Qy  2641  GCAGCCGTTAATGAGATGCTGGCACGCGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT   2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2740  GCGGCAGTTAATGAGATGCTGGCACGCGCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT   2799

Qy  2701  GGCGACCTCGCGCACTCCACCGTGTGGTGCGGGTGTAGATCCAGACAGACTTTGCTGCGAGAC   2760
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2800  GGCGACCTCGCGCACTCCACCGTGTGGTGCGGGTGTGAGATCCAGACAGACTTTGCTGCGAT   2859

Qy  2761  CCACAAAGTACGACATCCCAGACTCTGTCATCGGCTTCCTGCGCGCCGAGCTTGGTAAC   2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2860  CCACAAAGTACGACATCCCAGACTCTGTCATCGGCTTCCTGCGCGCCGAGCTTGGTAAC   2919

Qy  2821  CCTCCAGGTGGCTGGCTGGACGGCACTGCGCACACTGCCGCACTGGAAGGCCGCTCCGAAGGC   2880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2920  CCTCCAGGTGGCTGGCTGGACGGCACTGGCGCACACTGCCGCACTGGAAGGCCGCTCCGAAGGC   2979

Qy  2881  AAGGCACCTCGACGGAAGTTCCTGAGGAAGAGCAGGCCGCACCTGCGCGCTGATGATTCC   2940
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2980  AAAGCACCTTTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTGCGCGCTGATGATTCC   3039

Qy  2941  AAGGAACGTCGCAAACAGCCTGTTCCCGGAAGCCAACCAACCGACGAAGAGTTCCTC   3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3040  AAGGAACGTCGCAACAGCCTCAACCGCCTGTTCCGGAAGCCAACTGAAGAGTTCCTC   3099

Qy  3001  GAGCACCGTGCGCGCTTCGGCAACACTCCGTCGATGATCGTGCAATTCTTCTACGGA   3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  3100  GAGCACCGTCGCCGCGTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTACGCG  3159
                ||||||||||||||||||||||||    |||||||||||||||||||||||||||||
Qy  3061  CTGGTCGAAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCCACCCCACTGCTTGTT  3120
          |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3160  CTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCCACCCCACTGCTTGTT   3219

Qy  3121  CGCCTGGATGCGATCTCTGAGCCAGCAGATAAGGGTATCGCGACGATAAGGGTATCGGCTAAGCGTC  3180
          ||||||||||| |||||||||||||||||||||||||||  ||||||||||||||||||
Db  3220  CGCCTGGATGCCATCTCTGAGCCAGCAGATAAGGGTATCGGCAATGTTGTGGCTAACGTC  3279

Qy  3181  AACGGCCAGATCCGCCCAATGCGCCCCAATGCGTGCGTGACCGTCCGTTGAGTCTGACCGGCAACC  3240
          |||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db  3280  AACGGCCAGATCCGCCCAATGCGCCCCAATGCGTGCGTGACCGTCCGTTGAGTCGTCACCGCAACC  3339

Qy  3241  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3340  GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC  3399

Qy  3301  ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3400  ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG  3459

Qy  3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGACGAAGATTGAACGCGTTGTG  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3460  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGACGAAAATCGATCGCGTTGTG  3519

Qy  3421  GTTCCTGCTGCTCAACGAAGGTGGCGACTTGAAGTGCCGACTTGATCGTCGTGTTCCTAA   3474
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3520  GTTCCTGCTGCTCAACGAAGGTGGCGACTTGAAGTGCCGACTTGATCGTCGTCATTCCTAA   3573
```

```
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGAAGCATCGATGAAATTATCGGTGCA  240

Qy   292  GCTAAAAAAGTAAAGCAGATGCCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCC  351
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   241  GCTAAAAAAGTAAAGCAGATGCCTATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC  300

Qy   352  CAGCTTGCCCGCGAGTGCGCGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTT   411
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   301  CAGCTTGCCCGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTT   360

Qy   412  CTTGATCTCACCGGTGATAAGTCTCGTGCGGTAACGCGCCGGAAGAAGGCTGGTCTGCCA  471
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   361  CTTGATCTCACCGGTGATAAGTCTCGTGCGGTAACCGCGCCGGAAGAAGGCTGGTCTGCCA  420

Qy   472  GTTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGC  531
          ||||||||||||||||||||||||||||||  ||||||||| ||||||||||||||||||
Db   421  GTTTTTGGCGGAATCCACCCCGAGCAAAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC  480

Qy   532  CAGACTTACCCCATCTTTGTAAAGGCAGTTGCCGGTGGCCAGTCGTATGCGCTTT   591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGCCAGTCGTATGCGCTTT   540

Qy   592  GTTTCTCACCTGATGAGCTTCCGCAAATTGCAACAGAAGCATCTCGTGAAGCTGAAGCG   651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTTGCTTCACCTGATGAGCTTCCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG   600

Qy   652  GCATTCGGCACGGTTCGGTATATGTCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA  711
          ||||||||||||||||  |||||||| ||||||||||||||||||||||||||||||||||
Db   601  GCTTTCGGGCGATGCGGGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCA  660

Qy   712  GTGCAGATCCTTGGCGATCGCACTGCAGTCGAGTTGTACACCTTTATGAACGTGACTGCTCA  771
          |||||||||||||| |||||||||||||||| ||||||||||||  ||||||||||||||||
Db   661  GTGCAGATCCTTGGCGATCACACTGGAGAGTTGTACACCTTTATGAACGTGACTGCTCTCA  720

Qy   772  CTGCAGGCGTGCTCACCAGCCAAAAAGTTGTCGAAATTGCGCACAGCAGATTTGGATCCAGAA  831
          |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTGCAGGCGTCGTCACCAGCCAAAAAGTTGTCGAAATTGCGCACAGCAGATTTGGATCCAGAA  780

Qy   832  CTGCGTGATCGGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCGATTGGTTACCAGGGGC  891
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CTGCGTGATCGGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCGATTGGTTACCAGGGGC  840

Qy   892  GCGGAACCGTGGAATTCTTGGTCTGCGATGAAAAGGGCAACCACGTTTCATCGAAATGAAC  951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GCGGAACCGTGGAATTCTTGGTCTGCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC  900

Qy   952  CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG  1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAG  960

Qy  1012  GCGCAGATGCGCTTGGCTGCAACCTTGACAACCTTGAAGAATTGGGTCTGACCCAAGATAAG  1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GCGCAGATGCGCTTGGCTGCAACCTTGACAACCTTGAAGAATTGGGTCTGACCCAAGATAAG  1020

Qy  1072  ATCAAGACCACCGGTGCAGACTGCAGTGCCGCGATCACCAGCGAAGATCCAAACAACGGC  1131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ATCAAGACCACCGGTGCAGACTGCAGTGCCGCGATCACCAGCGAAGATCCAAACAACGGC  1080

Qy  1132  TTCCGCCCAGATACCGGAACTATCACCGGCCTCACCAGCGGAGCTGGCGTTGCGT  1191
          |||||||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  1081  TTCCGCCCAGATACCGGAACTATCACCGGCCTCACCAGCGGAGCTGGCGTTGCGT  1140

Qy  1192  CTTGACGGTGCAGCTCAGCTCGGTGGCCACTCGAAACCGGTGAAATCACCGTGGCTG  1251
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  1141  CTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGGTGAAATCACCGTGGCTG  1200

Qy  1252  AAAATGACCTGCCGTGGTTCCGACTTTGAAGCTTGCTTCTGAAGCTGTCCGC  1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AAAATGACCTGCCGTGGTTCCGACTTTGAAGCTTGCTTCTGAAGCTGTCCGC  1260

Qy  1312  GCTGAGTTCACCGTGTCTGGTTGTTGCAACCAACATTGGTTTCTTGCGTGCTGCGG  1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 10

```
AX120849
LOCUS       AX120849                3420 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 765 from Patent EP1108790.
ACCESSION   AX120849
VERSION     AX120849.1  GI:14037564
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 3420)
  AUTHORS   Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
            Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
  TITLE     Novel polynucleotides
  JOURNAL   Patent: EP 1108790-A 765 20-JUN-2001;
            KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES             Location/Qualifiers
     source          1..3420
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
BASE COUNT      732 a   1004 c    939 g    745 t
ORIGIN
```

```
Query Match     96.3%;  Score 3344.8;  DB 6;  Length 3420;
Best Local Similarity   98.6%;  Pred. No. 0;
Matches 3373;  Conservative    0;  Mismatches   47;  Indels   0;  Gaps   0;

Qy    52  GTGTCGACTCACACATCTTCAAGCGTTCCGTCGCAGCATTCAAAAAGATCTTGGTAGCAAACCGC  111
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GTGTCGACTCACACATCTTCAAGCGTTCCGTCGCAGCATTCAAAAAGATCTTGGTAGCAAACCGC  60

Qy   112  GGGCAAATCGCCGTCCGTGCTTCCGTCGCGAGCACTCGAAACCGGTGAGACCACGGTAGCT  171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GGGCAAATCGCCGTCCGTGCTTCCGTCGCGAGCACTCGAAACCGGTGAGCACGGTAGCT  120

Qy   172  ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC  231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  ATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGC  180

Qy   232  ATTGGTACTGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCA  291
```

```
Db 1261 GCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGG 1320

Qy 1372 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTC 1431
        ||||||||||||||||||||||||||||||||||||| || || ||||||| |||||||
Db 1321 GAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTC 1380

Qy 1432 CTTCAGGCTCCACCTGCGATGATGAGCAGGACGCATCCTGGATTACTTGGCAGATGTC 1491
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 CTTCAGGCTCCACCTGCGATGATGAGCAGGACGCATCCTGGATTACTTGGCAGATGTC 1440

Qy 1492 ACCGTGAACAAGCCTCATGGTGTGCCTCCAAAGGATGTTGCAGCACCAATGCGATAAGCTG 1551
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 ACCGTGAACAAGCCTCATGGTGTGCCTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTG 1500

Qy 1552 CCCAACATCAAGGATCTGCCACTGCCACGCGGTTCCGTGACGCGCCTGAAGCAGCTTGGC 1611
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCGTGACCGCCTGAAGCAGCTTGGC 1560

Qy 1612 CCAGCGCGCGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1671
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 CCAGCGCGCGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620

Qy 1672 TTCCCGGATGCACACCAGTCTTGCTGCGACACCGAGTCCGCTCATTCGCACTGAAGCCT 1731
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 TTCCCGGATGCACACCAGTCTTTGCTGCGACACCGAGTCCGCTCATTCGCACTGAAGCCT 1680

Qy 1732 GCGGCAGAGGCCGTCGCAAAGCTGACTCCGGAGCTTTTGTCCGTGAGGCCTGGGGCGGC 1791
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 GCGGCAGAGGCCGTCGCAAAGCTGACTCCGGAGCTTTTGTCCGTGAGGCCTGGGGCGGC 1740

Qy 1792 GCGACCTACGATGTGGCGATGGCTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1851
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GCGACCTACGATGTGGCGATGGCTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800

Qy 1852 CTGCGCGAGGCGGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCCGCCAACACCGTG 1911
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 CTGCGCGAGGCGGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCCGCCAACACCGTG 1860

Qy 1912 GGATACACCCCGTACCCGGTCTCCGCGCGTTGTTGTTGAAGGAAGCTGCCAGCTCC 1971
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GGATACACCCCGTACCCGGTCTCCGCGCGTTGTTGTTGAAGGAAGCTGCCAGCTCC 1920

Qy 1972 GGGCGTGGACATCTTCCGGCATCTTCGACGCGGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2031
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 GGGCGTGGACATCTTCCGGCATCTTCGACGCGGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980

Qy 2032 ATCGACGAGCAGTCCTGGAGACCAACACCGCCGAGGTGGCTATGGCTTATTCTGGT 2091
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 ATCGACGAGCAGTCCTGGAGACCAACACCGCCGAGGTGGCTATGGCTTATTCTGGT 2040

Qy 2092 GATCTCTGATCCAAATGAAAAGCTCTACACCCCTGGATTACTACCTAAAGATGGCAGAG 2151
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 GATCTCTGATCCAAATGAAAAGCTCTACACCCCTGGATTACTACCTAAAGATGGCAGAG 2100

Qy 2152 GAGATCGTCAAGTCTCGGCGCTCACATTCTGGCCATTCTGGCCATTAAGGATATGCTGTCTGCTTCGC 2211
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 GAGATCGTCAAGTCTCGGCGCTCACATTCTGGCCATTCTGGCCATTAAGGATATGCTGTCTGCTTCGC 2160

Qy 2212 CCAGCTGCGGTAACCAAGCTGGTCATCACCGTCACTGCCGTGATTCGCACTGTCCAGTGCAC 2271
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 CCAGCTGCGGTAACCAAGCTGGTCATCACCGTCACTGCCGTGATTCGCACTGTCCAGTGCAC 2220

Qy 2272 GTGCACACCCACGACACTGCGTTGACGGTGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCT 2331
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 GTGCACACCCACGACACTGCGTTGACGGTGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCT 2280

Qy 2332 GGTGCAGATGCTCGTTGACGGTGCTTCCGCACCACTGTCTGCACACCTCCCAGCCATCC 2391
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 GGTGCAGATGCTCGTTGACGGTGCTTCCGCACCACTGTCTGCACACCTCCCAGCCATCC 2340

Qy 2392 CTGTCTGCCATTGTTGCTCGCGCGATACCGGTTCGCGCCGATACCCGTGCGCACACCGTTTGAGCCTCGAG 2451
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 CTGTCTGCCATTGTTGCTCGCGCGATACCGGTTCGCGCCGATACCCGTGCGCACACCGTTTGAGCCTCGAG 2400
```

```
Qy 2452 GCTGTTTCTGACCTTCGAGCCTGACTGGAGAAGCTGTGCGGCGGACTGTACCTGCCATTTGAG 2511
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 GCTGTTTCTGACCTTCGAGCCTGACTGGAGAAGCTGTGCGGCGGACTGTACCTGCCATTTGAG 2460

Qy 2512 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGGTCGCGTCTACCGGCCGAAATCCCGGACAGTTG 2571
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 TCTGGAACCCCAGGCCCAACCGGTCGCGTCTACCGGTCGCGTCTACCGGCCGAAATCCCGGACAGTTG 2520

Qy 2572 TCCAACCTGCGTGCACAGGCACCGGCCACTGGGCCTTGCTGCTGTCGATCGCTTCGATGCGTCATCGAA 2631
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 TCCAACCTGCGTGCACAGGCACCGGCCACTGGGCCTTGCTGCTGTCGATCGTTCGAACTCATCGAA 2580

Qy 2632 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGGCCAACCAAGGTCACCCCCATCCTCC 2691
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 GACAACTACGCAGCCGTTAATGAGATGCTGGGACGGCCAACCAAGGTCACCCCCATCCTCC 2640

Qy 2692 AAGGTTGTTGGGACCTGCACTCCACCTCCACCTGGTGTGCGGGTGTAGATCCAGAGACTTT 2751
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 AAGGTTGTTGGGACCTGCACTCCACCTCCACCTGGTGTGCGGGTGTGGATCCAGAGACTTT 2700

Qy 2752 GCTGCAGACCCACAAAAGTACGACACTCTGTCATCGCGTTCCTGCGCGGCGGAG 2811
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GCTGCCGATCCACAAAAGTACGACATCCAGACTCTGTCATCGCGTTCCTGCGCGGCGGAG 2760

Qy 2812 CTTGGTAACCCTCCAGTGGCTGGCCAGAACCACTGCGCCACCCGGCCACTGGAAGGCCGC 2871
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 CTTGGTAACCCTCCAGTGGCTGGCCAGAACCACTGCGCCACCCGGCCACTGGAAGGCCGC 2820

Qy 2872 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGACGAAGAGCAGGCTCACTCGACGCT 2931
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGACGAAGAGCAGGCTCACTCGACGCT 2880

Qy 2932 GATGATTCCAAGGAACGTCGCAAGAGCCTCAACCGGCTGTGTTCCCGAAGCCAACCGAA 2991
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 GATGATTCCAAGGAACGTCGCAATAGCCTCAACCGGCTGTGTTCCCGAAGCCAACCGAA 2940

Qy 2992 GAGTTCCTGAGCCGCTTCGGCGAGCTGGCTGCGCTGGATGATGATCGTGAATTC 3051
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 GAGTTCCTGAGCCGCTTCGGCGAGCTGGCTGCGCTGGATGATGATCGTGAATTC 3000

Qy 3052 TTTCTACGGACTGGTCGAGGGCGCCGGAGACTTTGATCGCGCTGCCACCCCA 3111
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TTTCTACGGACTGGTCGAGGGCGCCGGAGACTTTGATCGCGCTGCCACCCCA 3060

Qy 3112 CTGCTTGTTCGCCTGACTGCGGATGCGGCGTCTCTGAGCCGCAGACGATAAGGGTATGCGCAATGTGTG 3171
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 CTGCTTGTTCGCCTGACTGCGGATGCGGCGTCTCTGAGCCGCAGACGATAAGGGTATGCGCAATGTGTG 3120

Qy 3172 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGCGTGACCGCTCCGTTGAGTCTGTC 3231
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGCGTGACCGCTCCGTTGAGTCTGTC 3180

Qy 3232 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGCCATGTTGCTGCACCATTCGCT 3291
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGCCATGTTGCTGCACCATTCGCT 3240

Qy 3292 GGTGTTGTCACTGTTGCTGAACTGTTGCTGAAAGGTGATGAGGTCAAGGCTCGCAGTGCGCA 3351
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GGTGTTGTCACTGTTGCTGAACTGTTGCTGAAAGGTGATGAGGTCAAGGCTCGGAGTCGCA 3300

Qy 3352 ATCATCGAGGCTATGAGAGATGGAAGCAACAATCACTGCTTCTTCTTGTTGACGGCAAGATTGAA 3411
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 ATCATCGAGGCTATGAGAGATGGAAGCAACAATCACTGCTTCTTCTTGTTGACGGCAAATCGAT 3360

Qy 3412 CGCGTTGTGGTTCCTGCTGCGACGGACTTGATCGTGTCGTTTTCC 3471
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 CGCGTTGTGGTTCCTGCTGCAACGAAGGTGGCGACTTGATCGTGTCGTCGTTTCC 3420
```

RESULT 11
AB083299
LOCUS   AB083299   4851 bp   DNA   linear   BCT 10-APR-2002

DEFINITION   Corynebacterium efficiens pyc gene for pyruvate carboxylase,
             complete cds.
ACCESSION   AB083299
VERSION   AB083299.1 GI:20126703
KEYWORDS   .
SOURCE   Corynebacterium efficiens DNA.
  ORGANISM   Corynebacterium efficiens
             Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
             Actinomycetales; Corynebacterineae; Corynebacteriaceae;
             Corynebacterium.
REFERENCE   1
  AUTHORS   Akiyoshi,N., Nonaka,G., Kimura,E., Kawahara,Y. and Sugimoto,S.
  TITLE   Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete
             CDS
  JOURNAL   Published Only in Database (2002)
REFERENCE   2 (bases 1 to 4851)
  AUTHORS   Nonaka,G., Akiyoshi,N., Kimura,E., Kawahara,Y. and Sugimoto,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (08-APR-2002) Gen Nonaka, AJINOMOTO CO., INC.,
             Fermentation&Biotechnology Laboratories; 1-1 Suzuki-cho, Kawasaki,
             Kanagawa 210-8681, Japan (E-mail:gen_nonaka@ajinomoto.com,
             Tel:81-44-244-7138, Fax:81-44-244-4757)
FEATURES   Location/Qualifiers
     source   1..4851
                     /organism="Corynebacterium efficiens"
                     /db_xref="taxon:152794"
     gene   1157..4576
                     /gene="pyc"
     CDS   1157..4576
                     /gene="pyc"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="BAB88903.1"
                     /db_xref="GI:20126704"
                     /translation="MVTTTPSTLPAFKKILVANRGEIAVRAFRAAYETGAATVAIYPR
                     EDRGSFHRSFASEAVRIGTEGSPVKAYLDIDEIINAAKKVKADAVYPGYGFLSSNAQL
                     ARECAENGITFIGPTPEVLDLTGDKSKAVSAAKKAGLPVLAESTPSTDIEIVKSAEG
                     QTYPIFVKAVAGGGGRGMRFVEKPEDLRELARAEASREAEAAFGDGSVYVERAVIKPQH
                     IEVQILGDHTGDVIHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKFPCKSI
                     GYQGAGTVEFLVDEAGNHVFIEMNPRIQVEHTVTEEVTSVDLVKAQMHLAAGATLKEL
                     GLTQDKITTHGAALQCRITTEDPSNNFRPDTGVITAYRSPGGAGVRLDGAAQLGGEIT
                     AHFDSMLVKMTCRGSDFETAVSRAQRALAEFNVSGVATNIGFLRALLREEDFTKRRID
                     TGFIGSHQHLLQAPPADDREQGRILEYLADVTVNKPHGERPETARPIEKLPEVENIPLP
                     RGSRDRLKQLGPEGFARDLREQDALAVTDTTPRDAHQSLLATRVRSFALTPAARAVAK
                     LTPELLSVEAWGGATYDVAMRFLFEDPWARLDELREAMPNVIOMLLRGRNTVGYTPY
                     PDSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPAIDAVLETGTSVAEVAMAYSGDLS
                     NPGEKLYTLDYLNLAEQIVDSGAHILAIKDMAGLLRRAAAPKLVTALRREFDLPVHV
                     HTHDTRAGGQLATYLAAANAGADAVDAASAPLSGTTSQPSMSALVAAFAHTRRDTGLNL
                     QAVSDLEPYWEAVRGLYLPFESGTPGPTGRVVRHEIPGGQLSNLRAQAVALGLADRFE
                     LIEDYAAVANEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDFAADPQKYDIPDSVIA
                     FLRGELGTPPGGWPEPLRTRALEGRSQGKAPLAEIPAEEQAHLDSDDSAERRGTLNRL
                     LFPKPTEEFLEHRRRPGNTSALDDREFFYGLKEGREELIRLTGVSTPMVVRLDAVSEP
                     DDKGMRNVVVNNGQIRPIKVRDRSVESVTATAEKADATNKVHVRAAPFAGVVTVTVAE
                     GDEIKAGDAVAIIEAMKMEATITAPVDGVIDRVVVPAATKVEGGDLIVVVS"
BASE COUNT   851 a   1653 c   1500 g   847 t
ORIGIN

Query Match   64.9%;   Score 2256.2;   DB 1;   Length 4851;
Best Local Similarity   78.7%;   Pred. No. 0;
Matches 2734;   Conservative 0; Mismatches 733;   Indels   8;   Gaps   3;

Qy   1 GTGACTGCTATCACCC--TTGGCGGTCTCTTGTTGAAAGGAATAATTACTCAGTGTCGA   58
          ||||||||||||||||| ||| || |||||||| ||||| |||||| |||| || |||
Db 1107 GTGACTGCCATCACCCAACAGGCTGCCTCCTGCTGCGAGCACTGACGTGGTTTACAA 1166

Qy   59 CTCACACATCTTCAAGCGTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAA   118
          | | ||| |||| |||||| || |||||||||||||| || |||||||| |||
Db 1167 C---AACACCCTCCACGCTGCGCGGCCGTCAAACAGGCGTGGCCACCGAGGTGAAA 1223

Qy   119 TCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACC   178
          |||||| ||||| || ||||||| ||||||||| || || |||| |||||| || |||
Db 1224 TCGCGGTTGCGAGCATTCCGCCGCCTACGAGACCGGGCCGGCCCAACCGTGGCCATCTACC 1283

Qy   239 CTGAAGGCTCACCAGTCAAGCGTACTTGGACATCGATGAGAATTATCGGTGCAGCTAAAAA   298
          || || || ||||| ||||| || ||| ||| ||||||||||| ||| || ||||| |||
Db 1344 CCGAGGGCTCACCCGTCAACGCGTACCTCGATATTGATGATCATCAACGCCGCCAAGA 1403

Qy   299 AAGTTAAAGCAGCATGCTATTTACCCGGACGATATGGCTTCCTGTCTGAAAAATGCCAGCTTG   358
          ||||| || ||||||| || |||||||| |||||| || |||| ||||| |||||| ||||
Db 1404 AGGTGAAAGCGGACGCGGTCACCGGGTATGGTTTCCTTTCGGAAAATGCCCAGCTCG 1463

Qy   359 CCCGGCAGTGGCGGGAAAAACGGCATTACTTTTTATTGGCCCAACCCCAGAGGTTCTTGATC   418
          |||||| ||||| |||||||| ||||| ||||||| ||| ||||||||||||| ||||| |
Db 1464 CGCGTGAATGGCGCGGAGACGGCATTACCTTCATCGGTCCGACGCCCGAGGTGCTCGACC 1523

Qy   419 TCACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGGAGAAGGCTGGTCTGCCAGTTTTGG   478
          || ||||| ||||| |||||||| ||| ||||| |||||||| ||||| ||| ||| |||
Db 1524 TCACGGGCGACAAGTCCAAGGCTGTGTCCGCCGCGGAGAAGGCCGGGCTGCCGGTGCTGG 1583

Qy   479 CGGAATCCACCCCGAGCAAAAAACATCGATGACATCGATGAGATCGTCAAGAGTCGCAGACTT   538
          ||||||| ||||||||||||||| ||| |||||||||||||| |||||||||||| ||||||
Db 1584 CGGAATCCACCCCCAGCACCGACATCGAGATCGTGAAGAGCGCCGAGGGCCAGACCT 1643

Qy   539 ACCCCATCTTTGTAAAGGCAGTTGCCGGTGGCGGGGACGGCGTATGCGCTTTGTTTCTT   598
          |||||||||| |||||||||| || ||||| ||||| |||| ||||| |||| |||| ||
Db 1644 ACCCGATCTTCGTCAAGGCCGTCGCCGGTGGTGGTGGTCGGATGCGGTTCGTCGAGA 1703

Qy   599 CACCTGATGAGCTCCGCAAATTGGCAACAGAGCAATCTCGTGAAGCTGGAGCGGGCATTCG   658
          ||| ||||||| || ||||||||| || |||||| |||||| |||| |||||| || |||
Db 1704 AGCCCGAGGACCTGCGTGCCCTGGCCGCAGGGAGCCGTGATCAAATCCGCAGCAGTGCCGA 1763

Qy   659 GCGACGGTTCGTGTATATGTGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGA   718
          ||||||| || || ||||||| |||| |||||| ||| ||||||| || |||| ||||||
Db 1764 GTGACGATCGCTACGGTGGCGGGCGGCCTCGACGGGCGTGATCAAATCATCGAGGTGCAGA 1823

Qy   719 TCCTTGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGC   778
          |||||||| || || |||| |||| ||||| || ||||||||||| ||| || ||||||
Db 1824 TCCTCGGTGATCACACCGGCGATGTCATCCACCTGTATGAGCGGGACTGTTCCCTGCAGC 1883

Qy   779 GTGTCACCCAAAAAGTGTCGAAATTGCCCAGCAGCATTTGGATCCAGAACTGCCTG   838
          || || ||||| ||| || ||||||||| ||||| ||||| |||| ||| || |||
Db 1884 GCGCACCACCAGAAGGTCGTGGAGATCGCACCTGCCCAGCACCTGGACCCCGAGCTGCGCG 1943

Qy   839 ATCGCATTTGTCGGATGCAGTCAAGGTCGTGCCGCCTCCATTGGTTACCAGGGCGGGGAA   898
          ||||| ||||||| || ||||| ||| ||| || |||| |||||||||||| ||||| ||
Db 1944 ACCGCATCTGTGCCGATGCCGTGAAGTTCTGCAAATCCATCGGATACCAGGGCGCCGGCA 2003

Qy   899 GTGTGAAATTCTTGGTGCGATGAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTA   958
          ||| || |||| ||| || ||||| || ||||||||||| || ||| || ||||||
Db 2004 CCGTGGAGTTCCTCGTCGACGAGGCGGGCAACCACGTCTTCATTGAGATGAACCCCCGCA 2063

Qy   959 TCCAGGTTGAGCAGACCGTGACTGAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCAGA   1018
          ||||||||||| ||||| ||| || |||||||| ||||||| || ||||||||| |||||
Db 2064 TCCAGGTCGAGCAGACCGTGACCGAGGAGGTCACCTCCGTCGACCTGGTCAAGGCGCAGA 2123

Qy 1019 TGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGA   1078
          ||||| ||| || ||||| |||||| || |||||||| ||||||||| || |||||||||
Db 2124 TGCACCTGGCCGCAGGTGCCACCCTGAAGGAACTGGGCCTGACCCAGGACAAGATCACCA 2183

Qy 1079 CCCACGGTGCAGCTGCACTGCAGTGCCGCATCACGACGGAAGATCCAAACAACGGCTTCCGCC   1138
          |||| || ||||| ||| || || ||||||| || |||||||||| || || |||||| ||
Db 2184 CCCACGGCGCCGCCCTGCAGTGCCGCATCACCACGGAGGACCCGTCCAACAACTTCCGGC 2243

Qy 1139 CAGATACCGGAACTATCACCGCGTACCGCTCCAGGCGGCGAGTCGGCCTTGACGGGAAA   1198
          || ||| || || ||||| ||||| ||||| |||| ||| ||| || || ||||||||
Db 2244 CCGACACCGGTGTGATCACCGCGTACCGCTCCCCGGGTGGTGCGGGGTGTCCGACG 2303

Qy 1199 GTGCAGCTCAGCTCCGTGCGGCGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACC   1258
          ||| ||||| ||||| || || || ||||| ||||||||||| || ||| ||| |||||
Db 2304 GCGCAGCCCAGCTCGGCGGCGAGATCACCGCACATTTCGATTCCATGCTGGTCAAGATGA 2363

```
Db   3441 ATGCCGTCGAGCGCGCCGCCTCCGGTACCGGTTTGAGCCTCAACCGGTCGATGTCCG 3500
          ||||||     ||||||||||| ||| |||||||||||||||||| ||  ||||||||||

Qy   2399 CCATTGTTGCTGCATTCGGCGACCGTCGCGATACCGGTTTGAGCCTCAACTGCGTGTTT 2458
          |||||  || || |||| ||| || ||| ||||||||||||||||||||  ||||||| |
Db   3501 CTCTGGTTGCCGCGCCGTTGCGCACACCGGGCACACGGCTCAACCTGCAGGCCGTCT 3560

Qy   2459 CTGACCTTGCGAGCCGTACTGGAAGCTGTGCGGCGGACTGTACCTGCCATTGAGTCTGAA 2518
          |  ||| ||||   ||||||||||||||||| |||| |||||||||||||| ||||| |||
Db   3561 CCGACTTGGAACCGTACTGGGAGGGCCGGTCCGGCGGACTGTACCTGCCGTTTGAATCCGGCA 3620

Qy   2519 CCCCAGGCCCAACCGGTGCGTCTACCGGCCACGAAATCCCAGGCGGACAGTTGTCCAACC 2578
          |||| |||||||||||||| |||| |||||||||| |||||||  ||||| |||||||||||
Db   3621 CCCCGGGCCCGACCCGGTTTACCGGCCACGCCCGGCGGTCAGCTGTCTCCAACC 3680

Qy   2579 TGCGTGCACAGGCCACCGCACTGGCCTTGCTGATGCTTCATCGAAGACAACACT 2638
          |||||||||||||| ||||||||||||| ||||||||| ||||||||||||||||
Db   3681 TGCGTGCCCAGCCGCCGTGCACTGGCGTTGCTGCTCAGCTCATCGAGAGACTACT 3740

Qy   2639 ACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCAAGGTTG 2698
          ||||||||||||  |||||||| || ||||| || ||||||||||||||||||
Db   3741 ACGCGGCCGTCAACGCAGCTGGGGTCGTCCGACCAAGGTCACCCAAGGTTG 3800

Qy   2699 TTTGGGCGACCTGCACTCCACCGGTTGGTGGCGGGTGAGATCCAGCAGACTTTGCTGCAG 2758
          | |||||||||| ||||| |||||| || |||||||| |||||| ||| ||||| |||||
Db   3801 TCGGTGACCTGCACTGCACCGTCGTGCGGCGTGTGAGCGTGTGCGACGG 3860

Qy   2759 ACCCACAAAGTACGACATCCCAGACTCTGTTCATCGAAGGCTTGGTA 2818
          ||  ||||||| ||| |||||||| ||| |  ||||||||| |||||
Db   3861 ATCCGCAGAAGTACGACATCCCGATTCGGTCATCGGTCTCCTGGGTA 3920

Qy   2819 ACCCTCCAGGTGGCTGGCCGACAACCACTGCGACTGGAAGCCGCTCCAAG 2878
          || ||||||||| |||||| |||||||||| | |||| |||||||||||||
Db   3921 CCCCTCCCGGTGGCTGGCCGACAACCGCTCCGAGGGTCGCTCCCAGG 3980

Qy   2879 GCAAGGCACCTCGACGAAGTTCCTGAGGAAGAGCCGCACCTGACGCGTGATGATT 2938
          |||| |||     ||||  || ||| ||||| ||||| |||||||||| ||||||||
Db   3981 GTAAGGCCCGGGACGCGTGGGAGATCCCCGAGGAGCAGGGCCCCACCTTCGGATTCCGATGATT 4040

Qy   2939 CCAAGGAACGTGCAACAGCTCAACCGCTTCCCGGGTGTTCCCGGAAGAGTTCC 2998
          ||||||||||| |||||||||||||||||||   ||||||||||||||||||||
Db   4041 CCGCGGAGCCGGGCGTGCGGCACCCGCACCCCTGCGTGTTCCCGACCGAGGAGTTCC 4100

Qy   2999 TCGAGCACCGTGCTCGCCGCCGCTTCGGCGGAGACTTGATCGTGGATGATCGTGACG 3058
          |||| ||||||||| ||||| |||||||||||||| ||||||||||||||||| |||||
Db   4101 TTGAGCACCGTGTCGCCGCGCTTCGGCGGAGACTTCGGATCCGGAGTTCTTCTACG 4160

Qy   3059 GACTGGTGCAGGGCCCGGAGACTTTGATCTCCCAGCAATCGACG 3118
          |||| ||||||| |||| |||||||||||  |||| ||| |||||
Db   4161 GCTTGAAGGAGGAGACGTGAGCAGCTGAGCGTGTGTCCACCCGATGGTGG 4220

Qy   3119 TTTCGCCTGGATGCGATCTCTGAGCACGAGCAGAGCATAAGGGTATGGCCAACG 3178
          |||||| || ||||||| |||||||| |||||| ||| ||||| ||| |||||||
Db   4221 TCCGCCTGGATGCGGTGTCGGAACCGGATGCACAAAGGCATGCGGTTGGTCAACG 4280

Qy   3179 TCAACGGCGTGTCGCCGCATCGTCTCCGTTGAGTGTCGTCACCCGCAA 3238
          ||||||||| ||||||| ||||||||||||||| ||||||||||||||
Db   4281 TCAACGGCCAGATCGCCCGATCAAGGTGCGGAGGTGTCGTGGAGTCCGCCA 4340

Qy   3239 CCGCAGAAAGGCAGATTCCTCCAACAAGGGCCAGTCGTTGCTCGCCCAGCTG 3298
          |||| || || || ||||| |||||||| ||||||| ||||||| ||||||||
Db   4341 CCGGGGAGAGGCCGATGGGCCGATCAAGGTGGCCGGTCGTTCCGTGGTGGTGG 4400

Qy   3299 TCAAGTTGACTGTTGCTGCGCCGTCGGTCTGAAGCTGCACGTGCACA 3358
          ||||| |||| || |||| ||||| |||||||||| ||||||||||||
Db   4401 TCACTGACTGTTGCTGCGCCGTCGGATGCGTCATTG 4460

Qy   3359 CGGTAACGACACTGCGCGGGTGGCCAGTTGACTTGGTCTGCAGCTGACGCTGCACA 3418
          |||| ||||  |||||||||||||| |||  |||| ||||||||||||||| |||||
Db   4461 CGGCGCCAAACTGTGCCACCTGGCCTCTCGTGCCTGTCAGCTGTGTCGCAG 4520

Qy   3419 CCCAGACACTGCGGGTGGCCAGTTGGCTACTTTGCTGCAGCTCAGCTGTGCAG 3473
          |||||||| |   ||||||| ||||| ||||||||||||||||||||| |||||
          ||| || || || | |||| | || ||| ||||||||||||| |||| ||||| ||
```

```
Qy   1259 CCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGT 1318
          ||||||  ||||||||| ||| |  ||| || || ||||||||||| |||||||||||||
Db   2364 CCTGCCGCGCGGTTCCGGATTTCGAGACCGCGTGTTGCTGCTGGAGCGTGCCCTGGCGGAGT 2423

Qy   1319 TCACCGTCGTGGTGTGCAACAACATTGGTTTCTGGTGCGTTGCTTGTCTGGGGGAAGAGG 1378
          ||| ||| || || ||||||||| |  |||  || |||| |  ||||||||||||||||||
Db   2424 TCAACGTCTCCGGCGTGGCCACCGGTTGCTTCATCGGCTTCATCGGCGTTCCCACCACCTGA 2483

Qy   1379 ACTTCACTTCCAAGCGCCATCGCCACCGGATTTATCGGCGATCACCCACACCTTCCTTCAGG 1438
          |||||||  |||   ||||||||||| ||||| ||||||| |||||||||||||||||||||
Db   2484 ACTTCACCAAGAGGCGGCATCGCCACGGCTTCATCGGCGATCACCCACACCTGCTGCCAGG 2543

Qy   1439 CTCCACCTGCGATGATGAGCAGGAGGCATCCTGATTACTTGGCAGATGTCACCGTGA 1498
          ||||| ||||||||||||||||||||||||||| ||||||| |||||||||||||||||
Db   2544 CCCCACCTGCGATGATGAGCAGGCGGTTTACCGCGGTCAGCGTCTGCACCGTGA 2603

Qy   1499 ACAAGCCTCATGGTGCGTCCAAAGGATGTTGCAGCACCAATCGATGATAAGCTGCCCAACA 1558
          ||||||  ||| || ||| ||||| ||||| ||||||||| | ||||||||||||||||| |
Db   2604 ACAAACCCACGGTGAACGCCC---CGGATGGGCAGCCCGGTCCGATAGAGAAGCTGCCCGAGG 2660

Qy   1559 TCAAGGATCTGCCACTGCCAAAGCGACCGGTTCCCGTGACCGCTTGAAGCAGCTTGGCCCAGCCG 1618
          | ||||||||| || |||| ||||| ||| |||||||||| || |||||||||||| |||||||
Db   2661 TGGAGAACATCCCGCCGCTCCACGCGGTCCCGACCAGCAGTCGGAGCGTCGGCCCCCGGAGG 2720

Qy   1619 CGTTTGCTGCTCCGTGATCTCCGTGAGCAGGACCGACTGCAGTTACTGATACCACCTTCGCG 1678
          ||||| |||| ||| ||| |||||| || || |||||||||| |||||| ||||||||||||
Db   2721 GTTTCGCCCGCGGATCGGCGGCAACAGGATGCCCTGGCCGTCACCCGACACCTTCGCG 2780

Qy   1679 ATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGGCGGCAG 1738
          |||||| | |  | |||| |||| || ||| |||| ||| |||| || ||||||||| |||
Db   2781 ATGCCCACCCCTGCGTCCTGGTCGCACCCCGACCGCTCCTGGCTCATCGGTGCCCGGCGGCGGCGGC 2840

Qy   1739 AGGCCCTGCAAAGCTGACTCCTGAGCGTTTTGTCCGGGAGGCCTGGCCTCCAAGTCCGGACCT 1798
          || |||||||||||||||| |||||||| | |||||||||| || || ||||||||| |||||
Db   2841 GCGCCCTCGCAAAGCTGCAAAGTCTGGTGCGTGTCGGTGGAGGCCAAGTCCGGTGTGCCACCT 2900

Qy   1799 ACGATGTGGGCGATGCGTTCCTCTTTGAGATCCCGTGGGACAGGCTGCGACGAGCTGCGCG 1858
          ||||| ||||||| ||||| || |||| |||||||||||| |||| ||||||||| ||||||
Db   2901 ACGACGTGGGCGATGCGTGGCTTCTTCGAGGATCCGGCACGCGCCGGGGATGAGCTGCGTG 2960

Qy   1859 AGGCCGATGCCGAATGTAAAACATTCAGATGGCCGGCCGCCAACACCGTGGGATACA 1918
          || | ||||| ||||||||||| |||||||| |||||||| || |||||||||||||||
Db   2961 AGGCCGATGCCGAATGTGAACATCCAGATGGCCGTTCGGTGCCGTCGCGTCGGGTACA 3020

Qy   1919 CCCCGTACCCAGAGTCCGTCTGCGGCGGCGTTTGTTGTTAAGGAAGTGCGCAGCTGCGGGTGG 1978
          ||||||| | ||||| || || |||||| ||| || |||| ||||||| ||||||||||||||
Db   3021 CCCCGTACCCGGAGTCCGTCTGCGGCGGCGTTTGTGTGTCCAGGAGGAGCCGCAAGTCCGGTGG 3080

Qy   1979 ACATCTTCCGCCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACG 2038
          |||||||||| || ||||| ||||| |||| |||||| ||||||||| ||||| ||||||||
Db   3081 ACATCTTCCGCCATCTTCGACGCGCTCAACAAGGATCTGAGCAGCTGAGCATGCCCGCG 3140

Qy   2039 CAGTCCTGGAGACCAACCACCGCAGTCCGGAGGTGGCTATGGCTATTCTGTCGTGATCTCT 2098
          |||| ||||||||| |||||||| || ||| ||||| |  || ||| |||||||||| ||||
Db   3141 CCGTCCTGGAGACCGACCGGCGGCACCCGGAGGTGTCGGGAACCTGGCCGTACTCCGGTGACCTGT 3200

Qy   2099 CTGATCAAATGAAAAGCTCTACACCTGATTACTTAAAGATGGCAGAGAGGATCG 2158
          || || |||| ||||||  || |||||| |||| || ||| |||||||||||||
Db   3201 CCAATCGGGGGAGAAGCCGACACATCCTGACTACCTGGACTACTACCTGGCCAGCAGATCG 3260

Qy   2159 TCAAGTTCTGGCGCTCACCATTCGGCCATTCGGCTCGCCCAGCTGCACGTGCACA 2218
          | |||||   || || ||||||||| |||||| |||||||||||||||||| | ||
Db   3261 TCGACTTCGGTGCACACATCGGCGCACATCGGCTGCATCGGCGCCGGCTGCATGTCCACA 3320

Qy   2219 CGGTAACGAGCCTGCGGGTGGCCAGTTGCTTGGCTACTACTTTGCAGCTGACGCTGCACA 2278
          |||| || || ||||||||||||| |||  |||  |||||||||| || ||| ||||||||
Db   3321 CGGCGCCAAACTGGTACACGACATCTGGGCCTGAATTCGACCTGAATTCGACTCACGTGCACCA 3380

Qy   2279 CCCAGACACTGCGGGTGGCCAGTTGGCTACTACTTTGCTGCAGCTGTGTGCAG 2338
          |||| ||||||||| |||||||| ||||| |||||||| | ||||||| |||||
Db   3381 CCCAGACACGACACCGCGGGTGCAGCTGTCAGCTGGCACTTACTCGGCCCGCAACGCCGGGGGCCG 3440

Qy   2339 ATGCTGTTGACGGTGCTTCCGCACCCTGTCTGGCCACCACTTCCCAGCCACCTGTCTG 2398
```

```
Db   4521 TGGTGCCGCCGCCAAGGTCGAGGGCCGGCGACCTCATCGTGGTCGTCCTA 4575
```

RESULT 12
AX064993

```
LOCUS       AX064993              1719 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 119 from Patent WO0100844.
ACCESSION   AX064993
VERSION     AX064993.1  GI:12542705
KEYWORDS
SOURCE      Corynebacterium glutamicum.
ORGANISM    Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1 (bases 1 to 1719)
AUTHORS     Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberhauer,G.
TITLE       corynebacterium glutamicum genes encoding proteins involved in
            carbon metabolism and energy production
JOURNAL     Patent: WO 0100844-A 119 04-JAN-2001;
            BASF AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1..1719
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     CDS             101..1699
                     /note="RXN02328"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC25736.1"
                     /db_xref="GI:12542706"
                     /translation="MTAITLGGLLKGIITIVSTHTSSTLPAFKKILVANRGEIAVRA
                     FRAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYIDIDEIIGAAKKVK
                     ADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLA
                     ESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRKLATEASREAEAA
                     FGDGAVYVERAVINPQHIEVQIILGDHTGEVVHLYERDCSLQRRHQKVEIAPAQHLDP
                     ELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVD
                     LVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRSPG
                     GAGVRLDGAAQLGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIG
                     FLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADVTVNKPHGVRPK
                     DVAAPIDKLPNIKDLPLPRGSRDR"
BASE COUNT    405 a    457 c    468 g    389 t
ORIGIN
```

```
Query Match    45.0%;  Score 1564.8;  DB 6;  Length 1719;
Best Local Similarity  98.3%;
Matches 1592;  Conservative   0;  Mismatches   27;  Pred. No. 0;  Indels  1;  Gaps   1;
```

```
Qy      1 GTGACTGCTATCACCCTTGGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 GTGACTGCTATCACCCCTTGGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 160

Qy     61 CACACAATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 220

Qy    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 280

Qy    181 CGTGAAGATCGGGGGATCATTCCACCGGTACCTGGACATCGGATGAAATTATCGGTGCTACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 CGTGAAGATCGGGGGATCATTCCACCGGTACCTGGACATCGGATGAAATTATCGGTGCTACC 340

Qy    241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 400

Qy    301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 460

Qy    361 CGCGAGTGCGCGAAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGTTCTTGATCTC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    461 CGCGAGTGCGCGAAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGTTCTTGATCTC 520

Qy    421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGGCGAAAGAAGGCGTGGTCTGCCAGTTTTGGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGGCGAAAGAAGGCGTGGTCTGCCAGTTTTGGCG 580

Qy    481 GAATCCACCCCGAGCAAAACATCGATGAGATCGTTAAAAGCCGCTGAAGGCCAGACTTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581 GAATCCACCCCGAGCAAAACATCGATGAGATCGTTAAAAGCCGCTGAAGGCCAGACTTAC 640

Qy    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGCGGACCGGTATGCGCTTTGTTTCTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    641 CCCATCTTTGTGAAGGCAGTTGCCGGTGGCGGACCGGTATGCGCTTTGTTGTTCTTCA 700

Qy    601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGCCATTCGGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    701 CCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGCCTTTCGGC 760

Qy    661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    761 GATGGGCGCGTATATGTCGAACGTGCTGTTGATTAACCCTCAGCATATTGAAGTGCAGATC 820

Qy    721 CTTGGGCGATCCGACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    821 CTTGGGCGATCCGACACCGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 880

Qy    781 CGTCACCAAAAAGTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    881 CGTCACCAAAAAGTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 940

Qy    841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGGCGGGAACC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGGCGGGAACC 1000

Qy    901 GTGGAAATTCTTGGTCGTGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1001 GTGGAATTCTTGGTCGTCCATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 1060

Qy    961 CAGGTTGAGCACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1061 CAGGTTGAGCACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1120

Qy   1021 CGCTTGGCTGCCTGTCGTCAACCTTGAAGGAATTGGGTCGACCCAAGATAAGATCAAGACC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1121 CGCTTGGCTGCGTGGTCAACCTTGAAGGAATTGGGTCGACCCAAGATAAGATCAAGACC 1180

Qy   1081 CACGGTGCAGCACTGCAGTCGCCGTCCGCATCCGAAGATCCAAACAACACGGCTTCCGCCCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1181 CACGGTGCAGCACTGCAGTCGCCGTCCGCATCCGAAGATCCAAACAACACGGCTTCCGCCCA 1240

Qy   1141 GATACCGGAACTATCACCCGCTCACCAGGCGGACTGGCCGTTCGTCTTGACGGT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1241 GATACCGGAACTATCACCCGCTCACCAGGCGGACTGGCCGTTCGTCTTGACGGT 1300

Qy   1201 GCAGCTCAGCTCGGTCGGCGAAATCACCGGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1301 GCAGCTCAGCTCGGTCGGCGAAATCACCGGCACACTTTGACTCCATGCTGGTGAAAATGACC 1360

Qy   1261 TGCCGTGGTTCCGACTTTGAAACTGTCGTTGCTCGTGGCTCGTGACAGCGCGCGTTGGCTGAGTTC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1361 TGCCGTGGTTCCGACTTTGAAACTGTCGTTGCTCGTGGCTCGTGACAGCGCGCGTTGGCTGAGTTC 1420

Qy   1321 ACCGGTGTCTGGTGTTGCAACCAACATTGGTTCTTGCGTTCTTGCGCGGGAAGAGGAC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1421 ACCGGTGTCTGGTGTTGCAACCAACATTGGTTCTTGCGTTCTTGCGCGGGAAGAGGAC 1480

Qy   1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGCCACCGGATTCATTGCCGATCACCCGCTTGCC 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1481 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCTTGCC 1540

Qy   1441 CCACCTGCCGATGATGAGCAGCAGGGACGCCATCCTGACTTACTTGCAGGAGTTCTTGATCTC 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1541  CCACCTGCTGATGAGCAGGAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC  1600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1501  AAGCCTCATGGTGTGCGTCCACGCGGTTCCGTGACCGCGTCCGAAGGCTGGCCCAGCCG   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||     |||||
Db  1601  AAGCCTCATGGTGTGCGTGCGTCCACGCGGTTCCGTGACCGCGTCCTATCGATAAGCTGCCTAACATC  1660

Qy  1561  AAGGATCTGCCACTGCCACGCGGTTCCGTGACCGCGTCCGTGACCG-CTGAAGCCAGCTTGGCCCAGCCG   1620
          ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db  1661  AAGGATCTGCCACTGCCACGCGGTTCCGTGACCGCGTCCGTGACCG-CTGAAGCCAGCTTGGCCCAGCCG  1719

RESULT 13
AX064995
LOCUS       AX064995              1406 bp    DNA     linear    PAT 24-JAN-2001
DEFINITION  Sequence 121 from Patent WO00100844.
ACCESSION   AX064995
VERSION     AX064995.1  GI:12542707
KEYWORDS    .
SOURCE      Corynebacterium glutamicum.
  ORGANISM  Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE   1  (bases 1 to 1406)
  AUTHORS   Pompejus,M., Kroeger,B., Schroeder,H., Zelder,O. and Haberhauer,G.
  TITLE     Corynebacterium glutamicum genes encoding proteins involved in
            carbon metabolism and energy production
  JOURNAL   Patent: WO 0100844-A 121 04-JAN-2001;
            BASF AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1..1406
                     /organism="Corynebacterium glutamicum"
                     /db_xref="taxon:1718"
     CDS             <1..1386
                     /note="FRXA0328"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC25737.1"
                     /db_xref="GI:12542708"
                     /translation="ASEAVRIGTEGSPVKAYLDIDEIIGAAKKVKADAIYPGVGFLSE
                     NAQLARECAENGITFIGPTPEVLDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVK
                     SAEGTYPIFVKAVAGGGRGMRFVASPDELRKLATEASREABAAFGDGAVYVERAVI
                     NPQHIEVQLGDHTGEVVHLYERDCSLQRRHQKVVEIAPAQHLDPELRDRICADAVKF
                     CRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIQVEHTVTEEVTEVDLVKAQMRLAAGAT
                     LKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTTITAYRSPGGAGVRLDGAAQLG
                     GEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVATNIGFLRALLREEDFTS
                     KRIATGFIADHPHLLQAPPADDEQRILDYLADVTVNKPHGVRPKDVAAPIDKLPNIK
                     DLPLPRGSRDR"
BASE COUNT      325 a    381 c    385 g    315 t
ORIGIN
Query Match        38.9%;  Score 1351.8;  DB 6;  Length 1406;
Best Local Similarity   98.0%;  Pred. No. 1.3e-280;
Matches 1379;  Conservative     0;  Mismatches    27;  Indels    1;  Gaps    1;

Qy  214  GCTTCTGAAGCTGTCCGATTGGTACTGAAGGCGTCCACCAGTCAAGGCGTACCTGGACATC  273
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GCTTCTGAAGCTGTCCGATTGGTACCGAAGGCGTCCACCAGTCAAGGCGTACCTGGACATC   60

Qy  274  GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCCGGGATATGGC  333
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCCGGGATACGGC  120

Qy  334  TTCCTGTCTGAAAATGCCCAGCTTGCCCGAGTGCGCGAAAAACGGCATTACTTTTATT  393
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TTCCTGTCTGAAAATGCCCAGCTTGCCCGAGTGCGCGAAAAACGGCATTACTTTTATT  180

Qy  394  GGGCCAACCCCAGCTGTCCGATCCGTGATAAGTCTCGTGCGTTAACCGCCGCCG  453
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GGGCCCAACCCCAGCTGTCCGATCCGTGATAAGTCTCGCGCGGTTAACCGCCGCCG  240

Qy  454  AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC  513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  321  GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCGTGAC  1380
```

RESULT 13 alignment (main sequence block):

```
Db   241  AAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCGATGAGATC   300
Qy   514  GTTAAAAGCGGCTGAAGGCTGAAGGCAGCTTACCCCATCTTTGAAGGCAGTTGCCGTGGTGGC  573
          |||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db   301  GTTAAAAGCGGCTGAAGGCTGAAGGCAGCCAGACTTACCCCATCTTTGAAGGCAGTTGCCGTGGTGGC  360

Qy   574  GGACGCGGTATGCGTTCCGTGACCGCGTCCGAAGGCTCCGCAAATTGGCAACAGAAGCA  633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GGACGCGGTATGCGTTCCGTGACCGCGTCCGATGAAGCTTCCGCAAATTAGCAACAGAAGCA  420

Qy   634  TCTCGTGAAGTCGATGAAGCTGGGACGGTTCGGTATATGTCGAACGTGCTGTGATT  693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCTCGTGAAGTCGATGAAGCTGGCGGACGGTTCGGTATATGTCGAACGTGCTGTGATT  480

Qy   694  AACCCCCAGCACATTGAAGTGCAGATCCTTGGCGACTGGAGATCGCACTGGAGAAGTTGTACACCTT  753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AACCCCTCAGCATATTGAAGTGCAGATCCTTGGCGACTGGAGATCCACACTGGAGAAGTTGTACACCTT  540

Qy   754  TATGAACGTGACTGCTCACTGCAGCGTCGTCAGCCACCAAAAGTTGTCACCAAATTGCGCCAGCA  813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TATGAACGTGACTGCTCACTGCAGCGTCGTCAGCCACCAAAAGTTGTCGACCAAATTGCGCCAGCA  600

Qy   814  CAGCATTTGGATCCAGAACTGCGTGCATTGTGCGGATGCAGTCAGTCAGTTCTGCCGC  873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CAGCATTTGGATCCGAACTGCCGTGCATTGTGCGGATGCAGTCAGTCAGTTCTGCCGC  660

Qy   874  TCCATTGGTTACCAGGGCGCGGAACCGTGAATTCTTGGTGCGATGAAAAGGCGCAACCAC  933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TCCATTGGTTACCAGGGCGCGGAACCGTGGAATTCTTGGTGCGATGAAAAGGCGCAACCAC  720

Qy   934  GTTTTCATCGAAATGAACCCACGTATCCAGTTGAGCACTCCGGTGACCCGTGACTGAGAAGTCACC  993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GTCTTCATCGAAATGAACCCACGTATCCAGTTGAGCACTCCGGTGACCCGTGACTGAGAAGTCACC  780

Qy   994  GAGGTGGACCTGGTGAAGGCCAGATGCGCTTGGCTGCTGCTGCTGCAACCTTGAAGGAATTG  1053
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GAGGTGGACCTGGTGAAGGCCGAGATCGCGCTGGCTGCTGCTGCTGCAACCTTGAAGGAATTG  840

Qy  1054  GGTCTCGACCCAAGATAAGATCCAGTCAGTGCCGCATCACCACG  1113
          |||||||||||||||||||||||||||||||||||||||||||
Db   841  GGTCTCGACCCAAGATAAGATCCAGTCAGTGCCGCATCACCACG  900

Qy  1114  GAAGATCCAAACAACGGCTTCCGCAGATACCGGAACTATCACCGCGTACCGCTCACCA  1173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GAAGATCCAAACAACGGCTTCCGCAGATACCGGAACTATCACCGCGTACCGCTCACCA  960

Qy  1174  GGCGGAGCTGGCGTTCGTCTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGTCCGACTTGGTTCCGGAACATCACCGCACAC  1233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GGCGGAGCTGGCGTTCGTCTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGTCCGACTTGGTTCCGGAACATCACCGCACAC  1020

Qy  1234  TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGCGTTCCGACTTTGAAACTGCTGTGTGCT  1293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGCGTTCCGACTTTGAAACTGCTGTGTGCT  1080

Qy  1294  CGTGCACAGGCGGCGGTTGGCTGGCTGAGTTCACCGTGTCGTCGGTTCGTCGGTGTTGCGTCAACCAACATTGGTTTC  1353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CGTGCACAGGCGGCGGTTGGCTGGCTGAGTTCACCGTGTCGTCGGTTCGTCGGTGTTGCAACCAACATTGGTTTC  1140

Qy  1354  TTGCGTGCGGTTGCTGCGCGGGAAGAGGACTTCACTTCACAAGGCGCATCGGCCACCGGATTATC  1413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  TTGCGTGCGGTTGCTGCGCGGGAAGAGGACTTCACTTCACAAGGCGCATCGGCCACCGGATTCATT  1200

Qy  1414  GGCGGATCACCCACCTCCTTCAGGCTGCCGGATGATGAGCGAGGGACGCCATCCTG  1473
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCCGGATCACCCGCCACCTCCTTCAGGCTGCCGGATGATGAGCGAGGGACGGACGCCATCCTG  1260

Qy  1474  GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTCGTGCGTTCCACCGCCGCA  1533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTCGTGCGTTCCAAAGGATGTTGCA  1320

Qy  1534  GCACCAATCGATAAGCTGCCCAATCAAGGATCTGCCACCCCGAGCAAAAACATCGATGAGATC  1593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCGTGAC  1380
```

```
Qy   1594  CGCCTGAAGCAGCTTGGCCCAGCCGCG 1620
           || |||||||||||||||||||||||||
Db   1381  CG-CTGAAGCAGCTTGGCCCAGCCGCG 1406
```

```
RESULT 14
SCF11/c
LOCUS       SCF11              40104 bp    DNA     linear   BCT 12-MAY-2002
DEFINITION  Streptomyces coelicolor cosmid F11.
ACCESSION   AL132662
VERSION     AL132662.2  GI:20520901
KEYWORDS    3-oxoacyl-(acyl-carrier-protein) synthase; acyl carrier protein;
            acyltransferase; alpha-galactosidase; beta-hexosaminidase; cold
            shock protein; gluconolactonase precursor; histidine kinase
            protein; hydrolase; oxidoreductase; peptidase; pyruvate
            carboxylase; regulatory protein; response regulator; sugar
            hydrolase; sugar transporter; transcriptional regulator;
            transmembrane protein.
SOURCE      Streptomyces coelicolor A3(2).
ORGANISM    Streptomyces coelicolor A3(2)
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 40104)
  AUTHORS   Redenbach,M., Kieser,H.M., Denapaite,D., Eichner,A., Cullum,J.,
            Kinashi,H. and Hopwood,D.A.
  TITLE     A set of ordered cosmids and a detailed genetic and physical map
            for the 8 Mb Streptomyces coelicolor A3(2) chromosome
  JOURNAL   Mol. Microbiol. 21 (1), 77-96 (1996)
   MEDLINE   9700351
   PUBMED    8843436
REFERENCE   2  (bases 1 to 40104)
  AUTHORS   Seeger,K.J. and Harris,D.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 40104)
  AUTHORS   Cerdeno,A.M., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-1999) Streptomyces coelicolor sequencing project,
            Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
            CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
            David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
            Colney, Norwich, Norfolk NR4 7UH, UK
COMMENT     On May 9, 2002 this sequence version replaced gi:6137022.
            Notes:
            Streptomyces coelicolor sequencing at The Sanger Centre is funded
            by the BBSRC and Beowulf Genomics
            Details of S. coelicolor sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL; http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
            numbered using the following system eg SC7B7.01c. SC (S.
            coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
            strand).
            The more significant matches with motifs in the PROSITE database
            are also included but some of these may be fortuitous. The length
            in codons is given for each CDS.
            Usually the highest scoring match found by fasta -o is given for
            CDS which show significant similarity to other CDS in the database.
            The position of possible ribosome binding site sequences are given
            where these have been used to deduce the initiation codon. Gene
            prediction is based on positional base preference in codons using a
            specially developed Hidden Markov Model (Krogh et al., Nucleic
            Acids Research, 22(22):4768-4778(1994)) and the FramePlot program
            of Bibb et al., Gene 30:157-66(1984) as implemented at
            http://www.nih.go.jp/
            jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the
            correct initiation codon. Where possible we choose an initiation
            codon (atg, gtg, ttg or att)) which is preceded by an upstream
            ribosome binding site sequence (optimally 5-13bp before the
            initiation codon). If this cannot be identified we choose the most
            upstream initiation codon.
            IMPORTANT: This sequence MAY NOT be the entire insert of the
            sequenced clone. It may be shorter because we only sequence
            overlapping sections once, or longer, because we arrange for a
```

```
                    small overlap between neighbouring submissions. Cosmid F11 Overlaps
                    with cosmid F6 on the AseI-F genomic restriction fragment.
FEATURES             Location/Qualifiers
     source          1..40104
                     /organism="Streptomyces coelicolor A3(2)"
                     /strain="A3(2)"
                     /db_xref="taxon:100226"
                     /clone="cosmid F11"
     RBS             252..255
     gene            274..1065
                     /gene="SCF11.02"
                     /note="SCO0522"
     CDS             274..1065
                     /gene="SCF11.02"
                     /note="SCF11.02, probable oxidoreductase, len: 262 aa;
                     similar to TR:O54197 (EMBL:AJ000671) Streptomyces
                     clavuligerus clavulanate-9-aldehyde reductase, 247 aa;
                     fasta scores: opt: 501 z-score: 567.7 E(): 3e-24; 39.1%
                     identity in 238 aa overlap. Contains Pfam match to entry
                     PF00106 adh_short, short chain dehydrogenase and prosite
                     match to entry PS00061 Short-chain
                     dehydrogenases/reductases family signature, also similar
                     to Streptomyces coelicolor SCF6.15c, 245 aa; fasta scores:
                     opt: 818 z-score: 805.2 E():0; 56.5% identity in 237 aa
                     overlap"
                     /codon_start=1
                     /transl_table=11
                     /product="putative oxidoreductase"
                     /protein_id="CAB59579.1"
                     /db_xref="GI:6137024"
                     /db_xref="SPTREMBL:Q9RK87"
                     /translation="MSRTDSAQQMRTIHRKDIVMTQTSKVVLVTGASSGIGEATALRL
                     AADGHRVFLGARRTERLEKLAARIAEDGGTAGYRRLDVTDAADVRAFVSAAVERWGRL
                     DVIVNNAGVMPLSPLSELKVAEWDRMLDVNVRGVLHGIAAALPVMRAQGGHVVNIAS
                     VGAYEVVPTAAVYCATKFAVRALSEGLRQESAGDIRVSVVSPGVTESELADSISDPRA
                     REDMKTYRSVAVPASAIADAIFAVSRPAEVDVNEIVVRPAASAQ"
     misc_feature    346..897
                     /gene="SCF11.02"
                     /note="Pfam match to entry PF00106 adh_short, short chain
                     dehydrogenase, score 248.20, E-value 1.2e-70"
     misc_feature    751..837
                     /gene="SCF11.02"
                     /note="PS00061 Short-chain dehydrogenases/reductases
                     family signature"
     gene            complement(1093..1530)
                     /gene="SCF11.03c"
                     /note="SCO0523"
     CDS             complement(1093..1530)
                     /gene="SCF11.03c"
                     /note="SCF11.03c, probable transcriptional regulator, len:
                     145 aa; similar to SW:ASNC_ECOLI (EMBL:AE000451)
                     Escherichia coli, regulatory protein AsnC, 152 aa; fasta
                     scores: opt: 222 z-score: 287.1 E(): 1.3e-08; 26.9%
                     identity in 145 aa overlap. Contains Pfam match to entry
                     PF01037 ASNC_trans_reg, AsnC family"
                     /codon_start=1
                     /transl_table=11
                     /product="putative transcriptional regulator"
                     /protein_id="CAB59580.1"
                     /db_xref="GI:6137025"
                     /db_xref="SPTREMBL:Q9RK86"
                     /translation="MDEIDRRLIALLQQDATRPYAALGEVGLSAGAAHERVRKLREQ
                     RVIRSTTVDVDPEALDRGVLAFVLVDSTAWMGDRAEDFAAVPEIQEAHVIAGSAAVLV
                     KVRTATTERLQDVLRRLYAIDGVSGTQATVVLETFFERPLDAG"
     misc_feature    complement(1162..1458)
                     /gene="SCF11.03c"
                     /note="Pfam match to entry PF01037 ASNC_trans_reg, AsnC
                     family, score 52.20, E-value 1.1e-11"
     RBS             1606..1610
     gene            1615..2523
                     /gene="SCF11.04"
                     /note="SCO0524"
     CDS             1615..2523
```

```
          /gene="SCF11.04"
          /note="SCF11.04, possible gluconolactonase, len: 302 aa;
          similar to SW:GNL_ZYMMO (EMBL:X67189) Zymomonas mobilis
          gluconolactonase precursor (EC 3.1.1.17)
          D-glucono-delta-lactone lactonohydrolase; fasta scores:
          opt: 326 z-score: 392.6 E(): 1.7e-14; 28.7% identity in
          258 aa overlap"
          /codon_start=1
          /transl_table=11
          /product="possible gluconolactonase precursor"
          /protein_id="CAB59581.1"
          /db_xref="GI:6137026"
          /db_xref="SPTREMBL:Q9RK85"
          /translation="MSTDGPYEILDDRFRTGRCANGDNRLEVLIYDGCRWAEGPLYLPA
          WRQLWSDIPNDRILRWDEATGSVGVFRAPAGHSNGNTLDRQGRLVTCEQGNRRVTRT
          EPDGRVTVLAERFPDGRRLNSPNDSVVRSDGTIWFSDPDFGITSDYEGHRAESEIGACN
          VYRIDPVSGQVRLAADGFDGPNGVILTPDERRLFVSDSRAARIHAFDIREDGTLSDGK
          VFAAEGRDVHFDNIRFDDEGRLWAAALHDGVHCYPPDGTLIGRLRVPEPVSNVAFGGP
          KNNRLFITATTSLYSLVMSVTGAPRL"

gene      2654..3469
          /gene="SCF11.05"
          /note="SCO0525"

CDS       2654..3469
          /gene="SCF11.05"
          /note="SCF11.05, hypothetical protein, len: 271 aa;
          similar to various hypothetical proteins, e.g. TR:CAB45341
          (EMBL:AL079345) Streptomyces coelicolor SCE66.05C
          hypothetical 33.1 KD protein, 300 aa; fasta scores: opt:
          808 z-score: 960.4 E(): 0; 53.2% identity in 250 aa
          overlap"
          /codon_start=1
          /transl_table=11
          /product="conserved hypothetical protein SCF11.05"
          /protein_id="CAB59582.1"
          /db_xref="GI:6137027"
          /db_xref="SPTREMBL:Q9RK84"
          /translation="MSDPSTTPGHTAQQKIDTSVPHSARIWNYWLGGKDNYPVDEQAG
          DAYTAVFPGIVTVARSSRAFLRRNITHLVAEAGIRQFLDIGTGLPTADNTHEVAQRIA
          PESRIVYVDNDPMVLAHARALLYSTPEGATSYIDSDVLDPDTVLEAAARTLDFDRPIA
          LILSNILGHIPDYDQARSVAARLVDALPSGSYLSINDGSRGIDPEFDRAQDGYNESGA
          APYILRTVDQITAFFDGLELVEPGVVPVTRWRPDADTEDPGLIGEHGGLARKP"

RBS       3493..3497
gene      3507..4340
          /gene="SCF11.06"
          /note="SCO0526"

CDS       3507..4340
          /gene="SCF11.06"
          /note="SCF11.06, possible oxidoreductase, len: 277 aa;
          similar to TR:O53622 (EMBL:AL021428) Mycobacterium
          tuberculosis putative oxidoreductase MTV030.21c, 276 aa;
          fasta scores: opt: 276 z-score: 329.4 E(): 5.6e-11; 24.7%
          identity in 267 aa overlap and to SW:PRXC_PSEFL
          (EMBL:AF031153) Pseudomonas fluorescens non-heme
          chloroperodixase (EC 1.11.1.10) (chloride peroxidase)
          (CPO-F) (chloroperoxidase F) Cpo or CpoF; fasta scores:
          opt: 233 z-score: 279.3 E():3.5e-08; 26.1% identity in 283
          aa overlap. Contains Pfam match to entry PF00561
          abhydrolase, alpha/beta hydrolase fold"
          /codon_start=1
          /transl_table=11
          /product="possible oxidoreductase"
          /protein_id="CAB59583.1"


Query Match            38.9%;  Score 1351;  DB 1;  Length 40104;
Best Local Similarity  63.5%;  Pred. No. 1.7e-280;
Matches 2154;  Conservative    0;  Mismatches 1215;  Indels   24;  Gaps    5;
```

```
Db   32459 TACGAGCTCGGCGCGGCACCGTCGCCGTCTTCCCGCAGGAGGACCGCAATTCGCTGCAC 32400
                 |||||  || ||||||    |  |  |||  |    | |||  | | | |||| 
Qy     205 CGCTCTTTGCTTCTCGAAGCTGTCCGCATTGTACTGAAGGCTCACCAGTCAAGGCGTAC 264

Db   32399 CGGTCCAAGGCGCCAGAGGCCTACGAGATCGGGGAGAGCCAGGGGCATCCGTCCGGCGTAC 32340
                 |||  | || |   | |||||          |    ||    ||   
Qy     265 CTGGACAATGCGATGAATTATCGGTGCAGCTAAAAAGTAAAGCAGATGCTATTTACCCG 324

Db   32339 CTCTCCGTGGAGGAGATCGTGCGGCCCCGCCGGTCGCCGTGCCCGGAAAACGGCCGCTACCCG 32280
Qy     325 GGATATGGCTTCCGTGTCGAAAATGCCCAGCTTGCGCCGGAAAAACGGCATT 384

Db   32279 GGCTACGGCTTCCTGTCCGAGAACCCGACTCGCCCGCCGCCTGCTGGAGAGCCAGGGCGGATC 32220
Qy     385 ACTTTTATTGGCCAACCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCGTGCGTGGTA 444

Db   32219 ACCTTCGTCGGTCCCAGCGCCGGATCCGTGCGTGAGGAACTGGACACCGGCACGGGCCGTG 32160
Qy     445 ACCGCCGCGAAGAAGGCTGTCTGCCAGTTTGGCGAATCCACCCCGAGCAAAAACATC 504

Db   32159 GCGCCGCCCGCCGGAGTACCGGTACCCGTGCGTCGGCCTTCGGCCCCTCCACCGACGTG 32100
Qy     505 GATGACATCGTTAAAGCGCTGAAGGCGCCAGACTTACCCCATCTTTGTAAGGCAGTGCC 564

Db   32099 GACGAACTCGTACGGCGCCGCCGACGACGTCGGCGTTCCCGTGTTCGTCAAGGCGGTCGCG 32040
Qy     565 GGTGTGGCGGACGCCGTATCGCGTGCCGTCTTGTTTCTTCACTGAGAGCTCCGCAAATTGGCA 624

Db   32039 GGCGGCAGCGATCCCGTGCGTGAGGAACCCGCCCAGTGCGCGAGGCCATC 31980
Qy     625 ACAGAAGCACATCTCGTGAAGCGGCATTCGGCGACCCTGGACGACGGTTCGGTATATGTCGAACGT 684

Db   31979 GAGGGCCGCCCTCCGCCGGCGCCTTCGGCCATCTCCACGGTCTTCCTGGAGAAG 31920
Qy     685 GCTGTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTGGCGTCGAGAGTCGCACTGGAGAAGTT 744

Db   31919 GCGGTCGTCGAAACCCCGCCCATCGAGGTGCAGATCCTCCGCCGACGCGGAGGGCGACGTC 31860
Qy     745 GTACACCTTTATGAACGTGACTGCTGCTCACTGCAGCGTCGTCACCAAAAGTTGTCGAAATT 804

Db   31859 ATCCACCTCTTCGAGCCGGACTGCTCCGGTGCAGCCGCGCGCCACCAGAGAAGGTGATGCTG 31800
Qy     805 GCGGCCAGCACAGCATTGGATCCAGAACTGCGTGATCGCATTGTGCCGGATGCAGTAAAG 864

Db   31799 GCGCCCCGCCAACCTCGACCCGCCGGGGGCACCTGCGGGATCTGGCGCCACCGGTGAAC 31740
Qy     865 TTCTGCGCTCCATTGGTTACCAGGGCGCGGGGAATTCTTGGTCGATGAAAAG 924

Db   31739 TTCGCCCGGCAGATCGGCTACCGGCGCAACGCGGGCACCGGCACCGTTCCTCGTCGACCGGGAC 31680
Qy     925 GGCAACCACGGTTTCATCGAAATGAACCCACGTATCCAGGTGAGCACACCGTGACTGAA 984

Db   31679 GGCAACCACCGTTCGTTCATCGAGGATGAACCCGCCGGATCATCCAGGTCGAGCGGTCACCGGAG 31620
Qy     985 GAAGTGCACCGAGGTGACCTGGTGAAGGGCCAGAATGCGCTTGGCTGCTGGTGTCAACCTTG 1044

Db   31619 GAGGTCACCGAGCGTGACCTGCACCTGCACCTGCATCGCCAGCCCATCGAGCCCCAGAGACGCTG 31560
Qy    1045 AAGGAATTGGGTCTGACCCAAGATAAGATCAAGAGCCCACGGTGCGACTGCAGTGCCGC 1104

Db   31559 GCCGACCTCGGACTCGCCCGCCAGGAGAACATCACCCTGCGCGGTGCCCAGCTCCAGTGCCGC 31500
Qy    1105 ATCACCACCGGAAGATCCAAACAGGCTTCGCCCAGATACCGGAACTATCACCGCGTAC 1164

Db   31499 ATCACCACCGAGGAACCCGGCTTCCGCCCGGACACACCGGCGGCAGATCAGCGCCTAC 31440
Qy    1165 CGCTCACCAGGCGGAGCTGGCGTTCGTCTTGA---CGGTGCAGCTCAGCTCGGTGCGAA 1221
                  ||| ||| ||
Db   31439 CGTTCGCCGGCGAGCGGCGTTCCGGGCATCGCGAGATCCGTGCGTTTCGGGGCTACGGAG 31380

Qy    1222 ATCACCGCCACACTTTGACTCCATGCTGGTGAAAAATGACCTGCCGTGTTCCGACTTTGAA 1281
           |||||  ||  | ||| |  ||||| |||   |   ||  ||  |    |  ||  |||
Db   31379 ATCAGCGGCCGCACTTCGACTCGATGCTCGTCAAGCTCCTCGCCGGGACGGGACTCCAC 31320
```

```
Db  29180  GAGGGCGGCGATCTTCGTCCAACTCGCCTGA 29148

RESULT 15
AF262949
LOCUS       AF262949              4218 bp    DNA     linear   BCT 07-NOV-2000
DEFINITION  Mycobacterium smegmatis hypothetical integral membrane protein
            gene, partial cds; pyruvate carboxylase (pyc) gene, complete cds;
            and hypothetical methyltransferase gene, partial cds.
ACCESSION   AF262949
VERSION     AF262949.1  GI:11118647
KEYWORDS    .
SOURCE      Mycobacterium smegmatis.
  ORGANISM  Mycobacterium smegmatis
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 4218)
  AUTHORS   Mukhopadhyay,B. and Purwantini,E.
  TITLE     Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
            rapid purification, molecular and biochemical characterization and
            regulation of the cellular level
  JOURNAL   Biochim. Biophys. Acta 1475 (3), 191-206 (2000)
   MEDLINE   20374587
   PUBMED    10913817
REFERENCE   2  (bases 1 to 4218)
  AUTHORS   Mukhopadhyay,B. and Purwantini,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAY-2000) Microbiology, University of Illinois, 601
            South Goodwin Avenue, Urbana, IL 61801, USA
FEATURES             Location/Qualifiers
     source          1..4218
                     /organism="Mycobacterium smegmatis"
                     /db_xref="taxon:1772"
     CDS             <1..564
                     /note="ORF1; similar to putative integral membrane protein
                     of Mycobacterium leprae encoded by GenBank Accession
                     Number AL023635 and to hypothetical protein Rv2968c of
                     Mycobacterium tuberculosis strain H37Rv encoded by GenBank
                     Accession Number 283018"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical integral membrane protein"
                     /protein_id="AAG30410.1"
                     /db_xref="GI:11118648"
                     /translation="PSAVWVLIAGVLGLAASLTLTVEKIELLLINPDYVPSCSINPVLS
                     CGSVMVTWQASLFGFPNPLIGIVAFSVVLTGVLAVAGVRLPRNYWAGLATGTALGAV
                     FVHWLIFQSLYRIGALCPYCMVVWAVTIPLAVVTAVIALRARAGDADAGPGNIVHQWR
                     WSLVALWFTALVLILVRFWEYWSTLL"
     gene            600..3983
                     /gene="pyc"
     CDS             600..3983
                     /gene="pyc"
                     /EC_number="6.4.1.1"
                     /function="ATP-dependent pyruvate carboxylase"
                     /note="purified native enzyme shows ATP-dependent pyruvate
                     carboxylase activity in vitro; NH2-terminus of mature PYC
                     protein confirmed by Edman degradation"
                     /codon_start=1
                     /transl_table=11
                     /product="pyruvate carboxylase"
                     /protein_id="AAG30411.1"
                     /db_xref="GI:11118649"
                     /translation="MISKVLVANRGEIAIRAFRAAYEMGIATVAVYPYEDRNSLHRLK
                     ADESYQIGEVGHPVRAYLSVDEIIRVAKHSGADAVYPGYGFLSENPDLAAKCARAGIT
                     FVGPSAEVLQLTGNKARAIAAARAAGLPVLSSSEPSSVDELMAAAADMEFPLFVKAV
                     SGGGRGMRRVTDRESLAEAIEAASREASAFGDASVYLEQAVLNPRHIEVQIIADGA
                     GNVMHLFERDCSVQRRHQKVELAPAPNLSDELRQICADAVAFARQIGYSCAGTVEF
                     LLDERGHHVFIECNPRIQVEHTVTEEITDVDLVSSQLRIAAGETLADLGLSQDRLVVR
                     GAAMQCRITTEVPANGFRPDTGRITAYRSPGGAGIRLDGGTNLGAEISAHFDSMLVKL
                     TCGRDFSAAASRARRALEFRIRGVSTNIPFLQAVIDDPDFRAGRVT'SFIDDRPHL
                     LTSRSPADRGTRILNYLADITVNKPHGERPSTVYPQDKLPPLDLQAPPAGSKQRLVE"
```

```
                     LGPEGFAGWLRESKAVGVTDTTFRDAHQSLLATRVRTGLLMVAPYVARSMPQLLSIE
                     CWGGATYDVALRFLKEDPWERLAALRESVPNICLQMLLRGRNTVGTPYPELVTSAFV
                     EEAAATGIDIPRIFDALNNVSMRPAIDAVRETGSTIAEVAMCYTGDLSDPAENLYTL
                     DYYLKLAEQIVEAGAHVLAIKDMAGLLRAPAAHTLVSALRSRFDLPVHVHTHDTPGGQ
                     LATYLAAWSAGADAVDGASAPMAGTTSQPALSSIVAAAAHTQYDTGLDLRAVCDLEPY
                     WEAVRKVYAPFESGLPGPTGVYTHEIPGGQLSNLRQQAIALGLGDRFEEIEANYAAA
                     DRVLGRLVKVTPSSKVVGDLALALVGAGITABEFAEDPAKYDIPDSVIGFLRGELGDP
                     PGGWPEPLRTKALQGRGPARPVEKLTADDEALLAQPGPKRQQALNRLLFPGTAEFEA
                     HRETYQDTSSLSANQFYGLRYGEEHRVQLERVQLEGVELLIGLEAISEADERGMRTVMCI1
                     NGQLRPVLVRDRSIASEVPRAEKADRNNADHIAAPFAGVTVGVAEGDSVDAGQTIAT
                     IEAMKMEAAITAPKAGTVARVAVAATAQVEGGDLLVVS"

     misc_feature    600..2159
                     /gene="pyc"
                     /note="Region: PYCA or ATP-binding domain"
     misc_feature    1455..1478
                     /gene="pyc"
                     /note="Region: PS00867 carbamoyl-phosphate synthase
                     subdomain signature 2"
     misc_feature    2160..3980
                     /gene="pyc"
                     /note="Region: PYCB or the biotin carrying domain"
     misc_feature    2679..2720
                     /gene="pyc"
                     /note="Region: PS00165 serine/threonine dehydratases
                     pyridoxal-phosphate attachment site"
     misc_feature    3867..3884
                     /gene="pyc"
                     /note="Region: PS00188 biotin-requiring enzymes attachment
                     site"
     CDS             3980..>4218
                     /note="ORF2; contains N-6 adenine-specific DNA methylase
                     signature, PS00092; similar to hypothetical protein
                     MLCB1243.09 of Mycobacterium leprae encoded by GenBank
                     Accession Number AL023635 and to hypothetical protein
                     Rv2966c of Mycobacterium tuberculosis strain H37Rv encoded
                     by GenBank Accession Number 283018"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical methyltransferase"
                     /protein_id="AAG30412.1"
                     /db_xref="GI:11118650"
                     /translation="MTRIIAGFLGGRRIEVPKSGTRPTSDRVREAVFNALAARLDFAG
                     LAVILDLYAGSGALGLEAISRGASSALFVESDQRAAA"

BASE COUNT      623 a    1433 c    1479 g    683 t
ORIGIN
Query Match      38.3%;  Score 1329.8;  DB 1;  Length 4218;
Best Local Similarity  62.6%;  Pred. No. 6.9e-276;
Matches 2123;  Conservative  0;  Mismatches 1257;  Indels  9;  Gaps  3;

Qy     86 TCAAAAAGATCTTGGTAGCAAACCGGCGGAAATCGGGGTCCGTGCTTCCGTGCAGCAC  145
          |||  ||| ||||  | | ||| |||||||||||| ||||||| |||||||||| |||||
Db    604 TCTCCAAGGTGCTCGTCGCCAACCGGTAGCTATTACCCCCGTGAAGATCGGGGATCATTCCACC  663

Qy    146 TCGAAACCGGTGCAGCCACGGTAGCTATTACCCCCGTGAAGATCGGGGATCATTCCACC  205
          |||| |||  ||||| |||||| ||| || || ||||||  |  ||| || |||| |||||
Db    664 ACGAGATGGGCCATCGGCCACGGTGGTCCGTGCTGGCGGTGATCCGTATCCGTACGGACCACCGGTATCGTCGTCCATC  723

Qy    206 GCTCTTTGCTTCTGAAGCTGCTCCGCATTGGTACTGGAGGCTCACCAGTCAAGGCGTACC  265
          |||  |||  ||||| | | ||||| ||| |||| ||||| ||  ||||| ||||| ||||
Db    724 GGCTCAAGGCCGACGAGTCATATCAGATCGGAGAGTCGGGTCATCCCGTCCGCGGTATC  783

Qy    266 TGGACATCGATGAAATTATCGGTGACGAGCTAAAAAGTAAAGCAGATGCTATTACCCGG  325
          |||||| ||||| || |||||||| ||||| |||||  |  ||||| |||||| ||||||
Db    784 TGTCGGTCGACGAGATCATCCGCGTCCAGCATTCGGCGCGCGACGCGGTGTACCCGG  843

Qy    326 GATATGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCCGAGTGCGCGGAAAACGCGATTA  385
          | || || ||| ||| | |||| ||||| || |||| |||||| |||| || |||||||
Db    844 GCTACCGCTTCCTGTCCGGAGAACCCCGATCGTCGCGGCCAAGTGCGCGGCCAAGTGCGCGCGAGGCGGGTATCA  903

Qy    386 CTTTTATTGGCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGTAA  445
          || | ||||| ||| ||  |||  || |||| |  ||||| || |  || ||||| ||||
Db    904 CGTTCGTGGGACCGTCGCCAGGTGCTGCAGCTCAACAAGGCACCGGGTAACAAGGCACCGCACCGCGGATCG  963
```

```
Qy    446  CCGCCGGAAGAGGCTGGTCTGCCCAGTTTTGGCGGAATCCACCCCGAGCAAAAACATCG   505
                |||||||||| ||| ||   |   |||||||  |||| | ||||| ||
Db    964  CCGCGGCGCGGGGGCCTTCCCGTGCGGAGCCGTCGTCGTCGTGG  1023

Qy    506  ATGACATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAAGGCAGTTGCCG   565
                  |||||||||||| ||||||| ||||| ||| ||||||||  ||
Db   1024  ACGAGTTGATGGCCGTGCCGCGACATGGAGTTCCCGCTGTTCGTCAAGGCGGTCTCGG  1083

Qy    566  GTCGTGGCGGACCCGGTATGCGCTTTGTTTCTTCACCTCGATGAGCTCCGCAAATTGGCAA  625
                |   ||  |||||||||||||
Db   1084  GTGGCGCGGGGGGGCATGCGCCGCGTCACCGACACCGCCTGCCGAGGCGATCG  1143

Qy    626  CAGAAGCATCTCGTGAAGCTGAAGGCTGCAATCGGCGACCGGTCGGTATATGTCGAACGTG  685
                   | ||||||| ||  | |||||||||||||||||
Db   1144  AGGCGGCCTCGCCGGGAGGCCGAGTCGGCCGTTCGGCGACGCGTCGGTGCCTGGACGCAGG  1203

Qy    686  CTGTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTGGCGATGCCACTGGAGAAGTTG   745
                ||| ||  |||||||||||||||   | |  ||||||||||
Db   1204  CCGTTGCTCAACCCGTCACATCGAGGTGCAGATCCTCGCCGACACGCGGGCAACGTCA  1263

Qy    746  TACACCTTTATGAAACGTGACTGCTCTCACTCGAGCGTCGTCACCAAAAAGTTGTCGAAATTG  805
                ||||  |  | || || ||||||||  ||||||||||
Db   1264  TGCACCTGTTCGAGCGTGACTGCAGCGTGCAGGCGCAGGTCGTCGAGCTGG  1323

Qy    806  CGGCAGCACAGCATTGGATCCCAGAACTGCGTGATCGGATCGATCGATGCAGTAAAGT   865
                ||||  |||||||||||||||   | |||||||||||||||
Db   1324  CGCCGGCCGCCCAACCTGAGTGACGAACTGCGCCAACAGATCTGCGGCGGTCACCTCGG  1383

Qy    866  TCTGCGCCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAAGG  925
                ||||||||||  ||| | |||   |  ||||||||||||||||
Db   1384  TCGCGGCCAGATCGGGTACTCGGTGCGCGGGGCACCGGTCGAGTTCTCGTCGACGAGCGCG  1443

Qy    926  GCAACCACGTTTTCATCGAAATGAAATGCTTGGTCGAGTATCCAGTTGAGCACCGTGACTGAAG  985
                |||||||||||  ||| |||||||||||||
Db   1444  GCCATCACGTGTTCATCGAGTCGAATCCGGCAATCCGGCACACCGGTGACCGAGG  1503

Qy    986  AAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGGTGCAACCTTGA  1045
                    |||||||||||||||  ||||||||||||||
Db   1504  AGATCACCGACGTGGACCTGGTCTCGTCCGCATCGCGCGAGACCGGCTCG  1563

Qy   1046  AGGAATTGGGTCTGACCCGAGCTGCTCCAAGATAAGATCAAGACCCACGGTGCAGCTCAGTGCCGCA  1105
                ||||||||  || |||||||||  | ||||||||||||
Db   1564  CCGATCTCGGTCTGTCCCAGGACCGGCTCGTGGTGCCGGTGCGTCGGCGCCACGGCGCTGGCGCA  1623

Qy   1106  TCACCACGGAAGATCCAAAACAGATCCAAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACC  1165
                || ||||||||||||||||||||||||||||||||
Db   1624  TCACCACCGAGGTCGTGTCGCCGCGCGCCCCAGCGCTTCCCGACACCGGGCCATCACCGCGTACC  1683

Qy   1166  GCTCCACCAGGCGGAGCTGGCGGGCATCGGGCGGCCGCCACCAACCTGGGTGCGGAGATCT  1225
                ||||||||||||  |||||||||| |||||||||||| ||||||||||||
Db   1684  GCTCGCGCCAGGCGGGCCGGGCGCGGGCATCGGGCGGCCGCCACCAACCTGGGTGCGGAGACCGG  1743

Qy   1226  CCGCACACTTGACTCATGCTGGTGAAAATGACCTGCCGCTGGTTCCGACTTTGAAACTG  1285
                ||||||  |||||||||||||||||||||||||||||
Db   1744  CGGATCTCGACTCGACTCCATGCTCAAGCTGGACGCGCGCCACTTCTCGGCCG  1803

Qy   1286  CTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTGCTCACCGTGTTGGTTGTTCGAACCAACA  1345
                |||||||||||||| |||||||||||||||||
Db   1804  CGGCCTCGCGCGGCGCGCCGCGCGCCCTGGCGAGTGCGGAGTGTGTCGACCAAACA  1863

Qy   1346  TTGGTTTCTTGCGGTGCGTGCTGCGGGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCG  1405
                |||||  |||||||| |   |||
Db   1864  TCCCGTTCCTCCAGGACGCGATCGACGATCCGGACTTCGCGCCGCGGACGGGTGACGACGT  1923

Qy   1406  GATTTATCGCGATCACCCACACCTCCCTTCAGGCTCCCACCTGGCTCGCCGATGATGAGCAGGGAC  1465
                |||||||||||||||||||
Db   1924  CGTTCATCGACGACCGGCCGGCCGCATCTATTGACCTCGGTGTCTCCGGTGTGTGCCTCCAAAGG  1983

Qy   1466  GCATCCTGATTACTTGCAGATGTCACCGTGAACAAGCCTCATGGTGCGTCCAAAGG  1525
                ||  ||| | |||||||||||||||||||||
Db   1984  GGATCCTCAACTACCTGCCGACATCACGTGGCGACATCACGCGGCGCAACGGCCTTCGA  2043

Qy   1526  ATGTTGCAGCCAATCGATAAGCTGCCCAACATCAAGGATCTGCCACTGCCACTGCCACCGCGGTT  1585
                |||                  ||||||||||||   | ||   || ||||| ||
Db   2044  CGGTTTACCCGCAGGACAAGCTGCCGCGCTGGGACGCGGCGGGCGGGCGCCGGGGGAT  2103

Qy   1586  CCGTGACCGCCTGAAGCAGTTGGCCCAGCGACGCGTTGCTCGTGTGATCTCCGTGCGGAATCCA  1645
                ||||||||||  | ||||||||||
Db   2104  CCAAACAGCGGCTCGTTGGAACTGGGGCCCGAGGGTTTCGGCGGCTGGCTGCGGAATCCA  2163

Qy   1646  ACGCACTGGCAGTTACTGATACCACCTTCCGGATGCAACCAGTCTTTGCTTGCTTGGCACCC  1705
                ||| ||||||   |||||||||||||||
Db   2164  AGGCCCGTCGGCGTCACCGACACGGCGGCCACCAGTCGCGCTGCTGGCCACGC  2223

Qy   1706  GAGTCCGCTCATTCGCACTGAAGCCTGCGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGC  1765
                    ||| |||||||||||||||||||||
Db   2224  GTGTGCGCCACCACCGGTCTGCTGATGGTGGCCGCCGTACGTCGCACGCTCGCCGCGCAGT  2283

Qy   1766  TTTTGTCCGTGGAGGCCTGGGCGCGGCGGGCGGCGACCTCGGCGATGGCGATGGCGATGGCGGTTCCTCTTTG  1825
                |||||||||||||||   ||| |||||||||||
Db   2284  TGCTGTGGATCGAGTGCTGGCTGGGGCCCTTCGCTTCGTCCTGAAGG  2343

Qy   1826  AGGATCCGTGGGACAGGTCGACGAGCAGCTGCGCGGAGGCGATGCCGAATGTAAACATTCAGA  1885
                || ||||||   |||| ||| ||||||||||||
Db   2344  AAGACCCGTGGGAGACGCGCTGGCCGGCGTCGCGGAGAGCGTGCCCAACATCTGCCTGCAGA  2403

Qy   1886  TGCTGCTTCGCGGCCGCAACACCGTGGATACACCCGTACCCAGACTCCGTCTGCCGCG  1945
                ||||||||||   |||  |||| ||||||||||||
Db   2404  TGCTGCTGCGGGGGACGCAACACCGTGGGCTACACCGCGTACCCGGGACTGGTCACCTCGG  2463

Qy   1946  CGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTCCGGGCGTCGGATCAATCTTCGCATCTTCGACACGGCCTTA  2005
                |||||||||||||||| ||||||||||||||
Db   2464  CGTTCGTGGAGGAGCCGCCGGACCGGTATCGCATCTTCCGGATCTTCGACATCTTGAGCGCTCA  2523

Qy   2006  ACGACCTCTCCCAGATGCGTCCAGCAATCGACAGTCAGTCTTGAGACACGCCAGTTCGTCGACCGCGGTAG  2065
                |||| ||||| |||| || |||| ||| |||||
Db   2524  ACAACGTCGAGTGCATGCGGCCCGATCGGGATCCCGGGAAACCGGTTCGACCATCG  2583

Qy   2066  CCGAGGTGGCTATGGCTTATTCTGGTGATCTCTGATCCCAAATGAAAAGCTCTACACCC  2125
                ||||||||||||||  || || |||||||||||||||||
Db   2584  CCGAAGTCGCGATCGTCTACAGCGGCGACCTCCAGCGAATCCCCGGGAGAACCTCTACACGC  2643

Qy   2126  TGGATTACTACCTAAAGATGGCGAGAGAGATCGTCAAGTCGTCAAGTCGTCGCGGCTCACATTCTCGGCCA  2185
                |||||||||||||||   ||||| ||||||  |||||||||||
Db   2644  TCGACTACTACCTGAAGCTGGCCGAGCAGATCGTCGAGGCGCCACGTGCTGGCGA  2703

Qy   2186  TTAAGGATATGTGGTCTGCTTGCTCTGCCTTCGCCGAGAATACGGAACTATCACCGCGTACC  2245
                |||| |||   |||||| || ||| ||||||||||||
Db   2704  TCAAGGACATGGCCGGTTCGTGTGCCGCGCCCCCCACACGCTCGTGAGCGCGGTTGC  2763

Qy   2246  GCCGTGAATTCGATCTGCCAGTGACGTGCACACCCACGACACTGGCGGGTGGCCAGTTGG  2305
                |||  ||||||  || | |||| |||| |||||
Db   2764  GCAGCCGGTTCGATCTGTCGATCTGCCGCAGTGCCGCCCCGGGCGCGGTCAGCTCG  2823

Qy   2306  CTACCTACTTTGCTGCCAGCTCAAGCTGGTGCAGATGCGTGTTGACGGTGCTTCCGACCAC  2365
                | |||||| |||||||||||||||||  |||||
Db   2824  CGACGTACCTCGCGGGCTCGGTCGTGCGGGCCGGACGGCCGGTGCGCGCCGGCCGA  2883

Qy   2366  TGTCTGGCACCACCTCCTGTCCAGCCATCCCTGTCTGCCATTGTTGCTGGCGGCCGCACACCC  2425
                || ||||| |||||| ||| || |||||||
Db   2884  TGGCCGGGACCACCAGCAGCCCCGCTCGATGTCGAGCTCGATCGTGGGCGGCCACACCC  2943

Qy   2426  GTCGCGGATACCGGTTGGGCCTCGCAGCGTGTTTCTGACCTCGAGCCTTGGGGAAGCTG  2485
                | |||||||||||||    ||||||||   ||||  ||||||||
Db   2944  AGTACGACACGGGCCTGGACCTGGCGTGCGTGACCCTACTGGGAGGCGG  3003

Qy   2486  TGCGCGGACTGTACCTGCCCATTTGAGTCTGAACCCCAGGGCCAGCCCGTCTACC  2545
                |||||||||||| ||| || |||| ||||  | | |||||||||||||
Db   3004  CGACGTACCTTCGCGGCCCGTTCGAGTCCGGCGATGAGCAGGGAC  3063

Qy   2546  GCCACGAAATCCCAGGCGGACAGTTGTCCAACCAACAGCCACCGGGCCACACCTGGGCC  2605
                ||||||  |||  |||||||   |   | ||||| ||||
Db   3064  CCCACGAGATTCCCGTGGGCAGTTGAGCAGCTTCAGCAGGCCACCCGGGGC  3123

Qy   2606  TTGCTGATCGCTTCGAGCTCATCGAAGACAACTACCGCAGCCGTTAATGAGATGCTGGGAC  2665
```

```
Db 3124 TCGGGCGACCGGTTCGAGGAGATCGAGGCCAATTACGCTGCGGCCGACCGGTTCTGGGAC 3183
        |||  ||  ||  |||||     ||| ||| |||         ||| |  ||||||||||

Qy 2666 GCCCAACCAAGTCACCCATCCTCCAAGGTTGTTGGCGACCTCGCACTCCACCTGGTTG 2725
        |||||     ||||   ||||   ||||||||||| ||  |  ||| |||||
Db 3184 GGCTCGTGAAGGTGACCCCGTCGTCGAAGGTGGTTCGGGGACCTGGCGCTCGGCGCTCGTGG 3243

Qy 2726 GTGCGGGGTGTAGATCCAGCAGACTTTGCTGCAGACCCACCAAAAGTACGACATCCCAGACT 2785
        |||||||| ||||||     ||| ||||| |||||    |||||||||||||||| ||||
Db 3244 GTGCGGGGCATCACCGCCGAGGAGTTCGCCGAGGATCCGCGAAGTACGACATCCCCGACA 3303

Qy 2786 CTGTCATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAACCAC 2845
        | ||||||||| ||||| |||||| |  ||   |||| || |||||||||| ||||||
Db 3304 GCGGTGATCGGCTTCCTGCGCGGCGGATCGGGGATCGGGGATGGCCGGAACCGT 3363

Qy 2846 TGCGCACCCGCGCACTGGAACGCCGTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTG 2905
        |||||        || ||||  ||||| ||||||||  ||  |||||| ||||| |||
Db 3364 TGCGCACCAAGCGCGTTCCAGGGCGCCGACCGGCCCCCGG---CCGGTCGAGAAGCTGACCG 3420

Qy 2906 AGGAAGAGCAGGCGCACCTGACGCCTGATGATTCCAAGGAACGTCCAACAGCCTCAACC 2965
        |         ||||| |    |||| ||||| |||| ||        | | ||| |||
Db 3421 CCGACCACGAGGCGGGTTGCTCGCCCAGCCCCGGCCCAAG---CGGCAGGCGCGTTGAACC 3477

Qy 2966 GCCTGCTGTTCCCGAAGCCAACGAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACA 3025
        ||||||||| |||||   ||| |  |||   |||| ||| ||  |  || ||||||||
Db 3478 GCCTGCTTTCCCCGGGCCCACCGGAGTTCGAGGCGCCACCGCGAAACCTACGGCGACA 3537

Qy 3026 CCTCTGCGCTGGATGATGTGAATTCTTCTACGGACTGGTCGAGGGCCGGAGACTTTGA 3085
        || ||| | ||| || |||| |  |||| ||||| ||||| |||| |||| ||||| |||
Db 3538 CCTCATCCCTCAGCGGCGAACCAGTTCTTCTACGGGCTGCGCTACGGCGAGGAGCACCGCG 3597

Qy 3086 TCCCGCTGCCAGATGTTGCGCACCCCACTGCTTGTTCGCCTCGGATCGGCATCTCTGAGCCAG 3145
        |       ||| |  || ||  ||  | |||  | |||  |||| || | |||||  ||||
Db 3598 TGCAACTCGGAACGTG---GCGTGGAACTGCGTGATCGGGCTTGAGGCGGCATCTCCGGAGGCCG 3654

Qy 3146 ACGATAAGGGTATGCGCAATGTTGTGCCAACGTCAACGGCCAGATCCGCCAATGCGTG 3205
        ||||   ||| | ||||| | | ||||||||| | || ||||||| |||||| ||| |
Db 3655 ACGAGGCGGGCATGCGGCATGGCGGTCATCATCAACGGTCAGCTGCGCCCGGTTCTCG 3714

Qy 3206 TGCGTGACCCGCTCCGTTGAGTCTTGTCACCCGAACCGCAGAAAAGGCCAGATTCCTCCAACA 3265
        | ||||| | |   |||  |||  ||||| | |||| |||||||  || |||  |||||||
Db 3715 TGCGGCACCGCAGCATCGCCAGCGAGGTGCCCGCCGAGGAGTGCCGGACCGCGACCGCAACAATG 3774

Qy 3266 AGGGCCCATGTTGCTGCACCATTCGCTGGTGTTGTCACTGTTGCTGACTGTTGCTGAAGGTGATG 3325
        |    |||||| ||  |  | | |||||| | || || ||  |||| | |||| |||| |  ||
Db 3775 CCGACCACATCGCCGCCGCCCTTCGCCCGGTGTGGTGACCGTCGGTGTCGCAGAAGGTGACT 3834

Qy 3326 AGGTCAAGGCTGGAGATGCAGTCGCAATCATCGCAATCATCGAGGCTATGAAAGATGGAAGCAACAATCA 3385
        ||   ||| ||| |||  |   |  |  |  || || |||||| || |||| ||| |||  ||||| |||
Db 3835 CGGTGACGCGGGACAAACCATCGCCGACGATCGAGCATCGAGGCGATGGAAGATGGAGGGCCGCCATCA 3894

Qy 3386 CTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTCGGTTCCTGCTGCAACGAAGGTGGAAG 3445
        |  |  | | |||| |  |||  |||||| |||  |||  | ||  | ||| | |||||| ||
Db 3895 CCGCGCCCAAGGCAGGCACCCTCGCGCGCGTCGCGCACGCGCCCAGGTCGAGG 3954

Qy 3446 GTGGGCACTTGATCGTCGTTCCTAA 3474
        |   || |||| | || ||  |
Db 3955 GCGGCGATCTGCTCGGTGGTGGTCAGCTGA 3983
```

Search completed: April 5, 2003, 10:09:08
Job time : 10481 secs
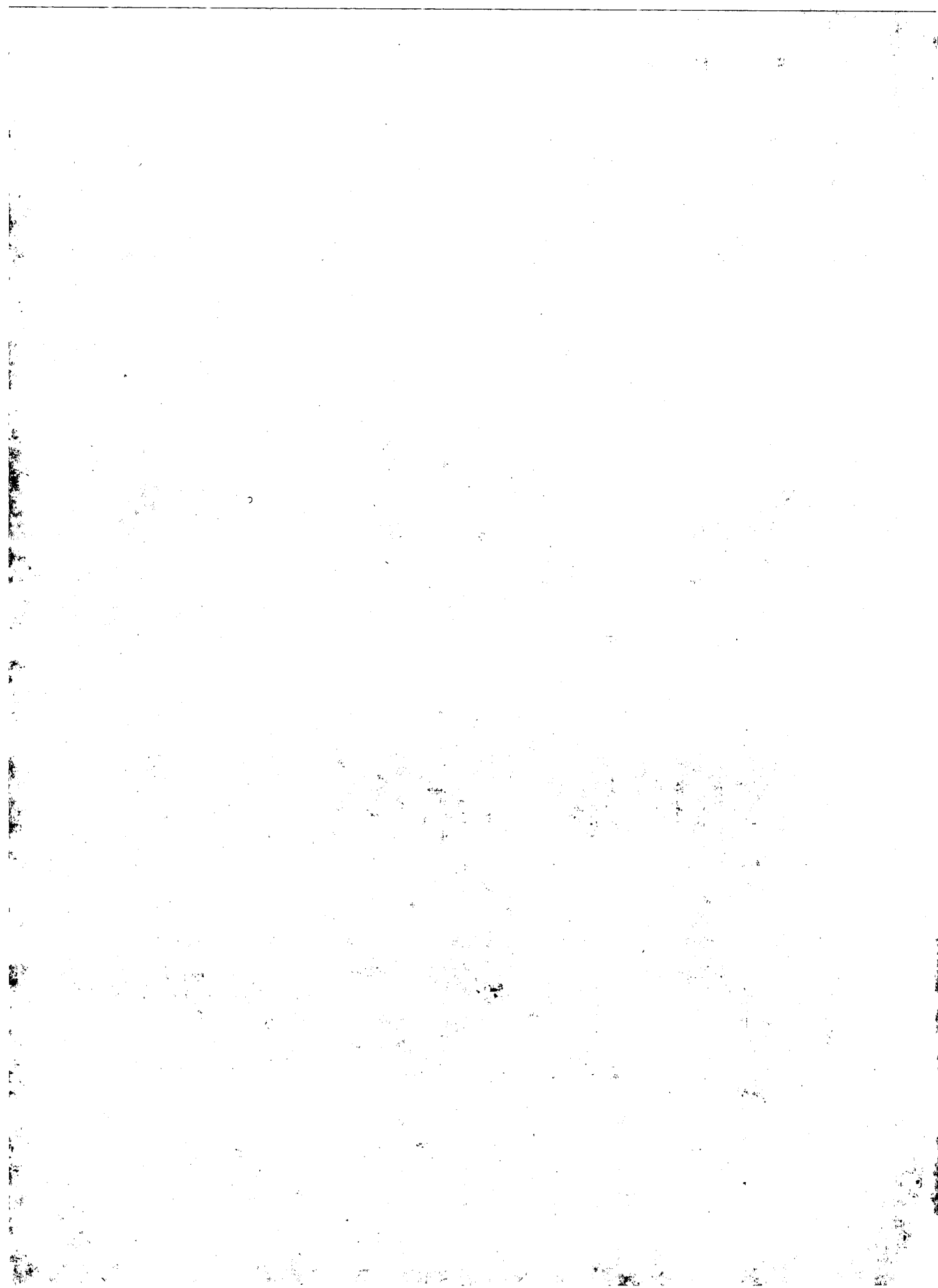
```
                    GenCore version 5.1.3
         Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       April   5, 2003, 04:30:25 ; Search time 762 Seconds
                                  (without alignments)
                          10266.984 Million cell updates/sec

Title:        US-09-974-973-1
Perfect score: 3474
Sequence:     1 gtcgactgctatcacccctgg.........tgatcgtcgtcgttcctaa 3474

Scoring table: IDENTITY.NUC
              Gapop 10.0 , Gapext 1.0

Searched:     2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters:       4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_101002:*
              1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
              2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
              3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
              4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
              5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
              6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
              7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
              8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
              9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
             10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
             11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
             12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
             13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
             14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
             15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
             16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
             17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
             18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
             19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
             20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
             21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
             22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
             23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
             24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

                        SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result          Query                                      Description
No.    Score   Match Length DB     ID
-------------------------------------------------------------------------
 1     3474    100.0  3474   24     ABK52832    Corynebacterium gl
 2     3398.8   97.8  3621   21     AAA47533    Sequence encoding
 3     3398.8   97.8  3621   22     AAF32165    Corynebacterium gl
 4     3398.8   97.8 349980  20     AAH68526    C glutamicum codin
 5     3397.2   97.8  3728   20     AAX24102    C. glutamicum pyru
 6     3344.8   96.3  3420   22     AAH65730    C glutamicum codin
 7     2256.2   64.0  4013   22     AAF87437    Corynebacterium th
 8     1564.8   45.0  1719   23     AAF71419    Corynebacterium gl
 9     1351.8   38.9  1406   22     AAF71420    Corynebacterium gl
```

```
10   1020.2   29.4   1083   22   AAF71417   Corynebacterium gl
11   1020.2   29.4   1083   22   AAF71418   Corynebacterium gl
12    883     25.4    939   22   AAF71415   Corynebacterium gl
13    883     25.4    939   22   AAF71416   Corynebacterium gl
14    681     19.6   3945   24   ABK63659   Rat sequence diffe
15    674.2   19.4   4017   24   ABN96532   Gene #3030 used to
16    665.2   19.1   4138   23   ABL19577   Drosophila melanog
17    665.2   19.1   4308   23   ABL19575   Drosophila melanog
18    665.2   19.1   4369   23   ABL02789   Drosophila melanog
19    620.2   17.9   4034   23   ABL20985   Drosophila melanog
20    606     17.4   4152   24   ABQ70879   Listeria monocytog
21    565.8   16.3 3011208  24   ABQ69245   Listeria innocua D
22    564     16.2   3462   22   AAS01509   Bacillus subtilis
23    502.8   14.5   3429   23   AAX13026   Enterococcus faeca
24    488.2   14.1  20072   20   AAS53072   Enterococcus faeca
25    486.8   14.0   3441   23   AAS51831   Staphylococcus aur
26    476.6   13.7   3456   24   AAH53527   S. epidermidis ope
27    476.6   13.7   3465   24   ABN91128   Staphylococcus epi
28    464.4   13.4   5030   18   AAV74635   Staphylococcus aur
29    428.8   12.3   3222   23   AAS54626   Staphylococcus aur
30    345.4    9.9   3432   22   AAH54232   S. epidermidis gen
31    308.2    8.9   7489   23   ABL20984   Drosophila melanog
32    308.2    8.9   7703   23   ABL19576   Drosophila melanog
33    308.2    8.9  12028   23   ABL02788   Drosophila melanog
34    308.2    8.9  12039   23   ABL19574   Drosophila melanog
35    305      8.8   1362   17   AAT43071   Synechococcus biot
36    305      8.8   1362   19   AAV33406   Synechococcus acet
37    303.4    8.7   1362   15   AAQ45330   fabG gene encoding
38    302.2    8.7   1350   23   AAS52596   E. coli DNA for ce
39    302.2    8.7   3077   16   AAQ79920   E. coli accB and a
40    302.2    8.7  14000   22   AAS46232   DNA encoding novel
41    297      8.5   3065   15   AAQ45328   fabG gene encoding
42    297      8.5   3065   17   AAT43070   Anabaena biotin ca
43    297      8.5   3065   19   AAV33405   Anabaena acetyl-Co
44    286      8.2   2238   24   ABK72711   Bacillus lichenifo
45    283.8    8.2   1350   23   AAS56150   Salmonella typhi D
```

ALIGNMENTS

```
RESULT 1
ABK52832
ID   ABK52832 standard; DNA; 3474 BP.
XX
AC   ABK52832;
XX
DT   27-AUG-2002   (first entry)
XX
DE   Corynebacterium glutamicum feedback-resistant pyruvate carboxylase gene.
XX
KW   Feedback-resistant; pyruvate carboxylase enzyme; gene; ds;
KW   aspartic acid feedback inhibition resistant.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3474
FT                   /*tag=  a
FT                   /product= "Corynebacterium glutamicum feedback-resistant
FT                              pyruvate carboxylase enzyme"
FT                   /transl_except= (pos:1..3; aa:Met)
XX
PN   WO200231158-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US31893.
XX
PR   13-OCT-2000; 2000US-239913P.
XX
PA   (ARCH ) ARCHER-DANIELS MIDLAND CO.
XX
```

```
PI   Hanke PD;
XX
DR   WPI; 2002-463267/49.
DR   P-PSDB; AAU98050.
XX
PT   Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT   polypeptide, useful for producing amino acids e.g. L-lysine,
PT   L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT   L-isoleucine
XX
PS   Claim 3; Fig 1; 42pp; English.
XX
CC   The present invention relates to a new mutated, feedback-resistant
CC   pyruvate carboxylase enzyme. The invention is useful for producing an
CC   amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC   by culturing a host cell in a suitable media and separating the amino
CC   acid from the medium. The vector of the invention is useful for
CC   replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC   resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC   replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC   a selectable marker gene through homologous recombination to form a first
CC   recombination strain, and replacing the selectable marker gene in the
CC   first recombinant strain, with feedback resistant pyruvate carboxylase
CC   gene through homologous recombination to form a second recombinant
CC   strain, where the homologous recombination in the above steps, occurs
CC   between the host cell and the vector. The feedback-resistant pyruvate
CC   carboxylase enzyme is resistant to feedback inhibition from aspartic
CC   acid. The present nucleic acid sequence encodes the feedback-resistant
CC   pyruvate carboxylase enzyme of the invention.
XX
SQ   Sequence 3474 BP; 749 A; 1013 C; 950 G; 762 T; 0 other;

Query Match             100.0%;  Score 3474;  DB 24;  Length 3474;
Best Local Similarity   100.0%;
Matches 3474;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;  Pred. No. 0;

Qy       1 GTGACTGCTATCACCCTTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTGACTGCTATCACCCTTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT    60

Qy      61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAAACCGGCGGAAATC   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAAACCGGCGGAAATC   120

Qy     121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCAGCTAGTTACCCC      180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCAGCTAGTTACCCC      180

Qy     181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCGAAGCTGTCCGCATTGGTACT   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCGAAGCTGTCCGCATTGGTACT   240

Qy     241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAGCATCGTCAGCTAAAAAA    300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAGCATCGTCAGCTAAAAAA    300

Qy     301 GTTAAAGCAGAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAAATGCCCAGCTTGCC   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GTTAAAGCAGAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAAATGCCCAGCTTGCC   360

Qy     361 CGCGAGTGCGCGGAAAAACGGCATTACTTTATTGGCCCAACCCCAACCCCAGAGGTTCTTGATCTC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CGCGAGTGCGCGGAAAAACGGCATTACTTTATTGGCCCAACCCCAACCCCAGAGGTTCTTGATCTC   420

Qy     421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCCGCGAAGAAGGCTGGTCTCCCAGTTTTGGCG    480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCCGCGAAGAAGGCTGGTCTCCCAGTTTTGGCG    480

Qy     481 GAATCCACCCCGAGCAAAAACATGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC    540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GAATCCACCCCGAGCAAAAACATGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC    540

Qy     541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGGTATGCGCGTTTGTTTCTTCA    600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGGTATGCGCGTTTGTTTCTTCA    600

Qy     601 CCTGATGAGCTCCGCAAATTGCAACAGAGCATCTCGTGAAGCTGAAGCGGCATTCGGC   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CCTGATGAGCTCCGCAAATTGCAACAGAGCATCTCGTGAAGCTGAAGCGGCATTCGGC   660

Qy     661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACAGCATTGAAGTGCAGATC   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACAGCATTGAAGTGCAGATC   720

Qy     721 CTTGGCGATCGCACTGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTTGGCGATCGCACTGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT   780

Qy     781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT   840

Qy     841 CGGATTTGTGCGGATGCAGTAAAAGTTCTGCCGCCTCCATTGGTTACCAGGGCGCGGGGAACC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CGCATTTGTGCGGATGCAGTAAAAGTTCTGCCGCCTCCATTGGTTACCAGGGCGCGGGGAACC   900

Qy     901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACCGTTTCATCGAAATGAACCCACGTATC   960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACCGTTTCATCGAAATGAACCCACGTATC   960

Qy     961 CAGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CAGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG   1020

Qy    1021 CGCTTGGCTGCTGCTGGTGTGCAACCTTGAAGGAATTGGGTCTGCTGCGACCCAAGATCAAGACC   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CGCTTGGCTGCTGCTGGTGTGCAACCTTGAAGGAATTGGGTCTGCTGCGACCCAAGATCAAGACC   1080

Qy    1081 CACGGTGCAGCACTGCAGTGCCGCATCACCGGAAGATCCAAACAACGGCTTCCGCCCA   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CACGGTGCAGCACTGCAGTGCCGCATCACCGGAAGATCCAAACAACGGCTTCCGCCCA   1140

Qy    1141 GATACCGGAACTATCACCGCGTCACCGGAGTCGGCGTTCGTCTTGACGGT   1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GATACCGGAACTATCACCGCGTCACCGGAGTCGGCGTTCGTCTTGACGGT   1200

Qy    1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGTGTGAAAATGACC   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGTGTGAAAATGACC   1260

Qy    1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTGTCGTCGTCAGCGCCGCGGTTGGCTGAGTTC   1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTGTCGTCGTCAGCGCCGCGGTTGGCTGAGTTC   1320

Qy    1321 ACCGTGTCCTGGTGTTGTCAACCAACATTGGTTTCTTGCGTGCGTTGCGTGCGTGGGGAAGAGAGAC   1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ACCGTGTCCTGGTGTTGTCAACCAACATTGGTTTCTTGCGTGCGTTGCGTGCGTGGGGAAGAGAGAC   1380

Qy    1381 TTCACTTCCAAGGCATCGCACCCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT   1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TTCACTTCCAAGGCATCGCACCCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT   1440

Qy    1441 CCACCTGCCGGATGAGAGCACGGGACACCCATCCTGGATTACTTGGCCAGATGTCACCGTGAAC   1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CCACCTGCCGGATGAGAGCACGGGACACCCATCCTGGATTACTTGGCCAGATGTCACCGTGAAC   1500

Qy    1501 AAGGCCATGGTGCGTCCAAAGGATGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC   1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AAGGCCATGGTGCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC   1560

Qy    1561 AAGGATCTGCCACTGCCACGCCGCCTGAAGGCTAAAAGGCTGAAGCCCAGCTTAC   1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AAGGATCTGCCACTGCCACGCCGCCTGGTTCCCGGTGACCGGCGTGAAGGCAGCTGAAGCCCAGCCCG   1620

Qy    1621 TTTGCTCGTGATCTCCGTGAGCAGGAGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT   1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy 2761 CCACAAAAGTACGACATCCCAGACTTCTGTCATCGGTCTTCCTGCGCGGCGAGCTTGGTAAC 2820
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 CCACAAAAGTACGACATCCCAGACTTCTGTCATCGGTCTTCCTGCGCGGCGAGCTTGGTAAC 2820

Qy 2821 CCTCCAGGTGGCTGGCAGAACCACTGCGCACCCGGCCACTGGAAGGCCGCTCCGAAGGC 2880
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 CCTCCAGGTGGCTGGCAGAACCACTGCGCACCCGGCCACTGGAAGGCCGCTCCGAAGGC 2880

Qy 2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGGACGCAGGCGCACCTGCGCTGATGATTCC 2940
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGGACGCAGGCGCACCTGCGCTGATGATTCC 2940

Qy 2941 AAGGAACGTCGCAACAGCCTCAACCGGCCTCAACAGCCTCAACCGGCCAACCGAAGAGTTCCTC 3000
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 AAGGAACGTCGCAACAGCCTCAACAGCCTCAACAGCCTGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000

Qy 3001 GAGCACGTCGCGCGTTCGGGCAACACACCTCTGCGCTGGCAACACCTGGATGATCGTGAATTCTTCTACGGA 3060
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 GAGCACGTCGCGCGTTCGGGCAACACACCTCTGCGCTGGCAACACCTGGATGATCGTGAATTCTTCTACGGA 3060

Qy 3061 CTGGTCGAGGGCCGGCAGACTTTGATCCGCCGGACATGTGCGCACCCCACTGCTTGTT 3120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 CTGGTCGAGGGCCGGCAGACTTTGATCCGCCGGACATGTGCGCACCCCACTGCTTGTT 3120

Qy 3121 CGCCTTGGATGCGATCTCTGAGCCAGCAGACGATAAGGTATGCCGCCAATGTTGTCGGCCAACGTC 3180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 CGCCTTGGATGCGATCTCTGAGCCAGCAGACGATAAGGTATGCCGCCAATGTTGTCGGCCAACGTC 3180

Qy 3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCCGTTCGTTGAGTTCGTGTCGTCACCGCCAACC 3240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCCGTTCGTTGAGTTCGTGTCGTCACCGCCAACC 3240

Qy 3241 GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCTGCGTGGTGTTGTC 3300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCTGCGTGGTGTTGTC 3300

Qy 3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGGCTGGAGATGCAGTGCAATCATCGAG 3360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGGCTGGAGATGCAGTGCAATCATCGAG 3360

Qy 3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAGATTGAACGCGTTGTG 3420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAGATTGAACGCGTTGTG 3420

Qy 3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGGCGACTTGATCGTCGTCGTTTCCTAA 3474
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGGCGACTTGATCGTCGTCGTTTCCTAA 3474
```

```
Db 1621 TTTGCTCGTGATCTCCGGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680

Qy 1681 GCCACCAGTCTTTGCTTGCCGACCCGAGTCGCTCATTCGGCACTGAAGCCTGCGGCGAGAG 1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 GCCACCAGTCTTTGCTTGCCGACCCGAGTCGCTCATTCGGCACTGAAGCCTGCGGCGAGAG 1740

Qy 1741 GCCGTCGCAAAGTCACTCCTGAGCTTTGTCCGTGGAGGCCGGCGGCGGCGACCTAC 1800
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GCCGTCGCAAAGTGACTCCTGAGCTTTGTCCGTGGAGGCCGGCGGCGGCGACCTAC 1800

Qy 1801 GATGTTGGCGATGCGTTTCCTCTTTGAGGATCGCAGCAGGCTGCACGAGGCTGCGCGAG 1860
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 GATGTTGGCGATGCGTTTCCTCTTTGAGGATCGCAGCAGGCTGCACGAGGCTGCGCGAG 1860

Qy 1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCGCCAACACCGTGGGATACACC 1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCGCCAACACCGTGGGATACACC 1920

Qy 1921 CCGTACCCAGACTCCGTCTGCCGCGGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 CCGTACCCAGACTCCGTCTGCCGCGGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980

Qy 1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040

Qy 2041 GTCCTGGAGACCAACACACCCGCGGAGTGGCTATGGCTTATTCTGGTGATCTTCT 2100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 GTCCTGGAGACCAACACACCCGCGGAGTGGCTATGGCTTATTCTGGTGATCTTCT 2100

Qy 2101 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCCTAAAGATGGCAGAGGAGATCGTC 2160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCCTAAAGATGGCAGAGGAGATCGTC 2160

Qy 2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTCGTGCTCGTGCTCCGCCCAGCTGCG 2220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTCGTGCTCGTGCTCCGCCCAGCTGCG 2220

Qy 2221 GTAACCAAGCTGGTCACCGACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 GTAACCAAGCTGGTCACCGACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280

Qy 2281 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTGCTGCAGCTCAAGGTGGTGCAGAT 2340
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTGCTGCAGCTCAAGGTGGTGCAGAT 2340

Qy 2341 GCTGTGTGACGGTGCTTCCCGGCCACCACTGCTCTCGGCCACCACTTCCCGACCATCCCGTCTGCC 2400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 GCTGTTGACGGTGCTTCCCGGCCACCACTGCTCTCGGCCACCACTTCCCGACCATCCCGTCTGCC 2400

Qy 2401 ATTGTTGCTGCAATTCGGCGACACCCGTCGCGCGACTACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 ATTGTTGCTGCAATTCGGCGACACCCGTCGCGCGACTACCGGTTTGAGCCTCGAGGCTGTTTCT 2460

Qy 2461 GACCTCGAGCCGTACTGGGAACGCTGTGCGCGGAGCTGTACCTGCCATTTGAGTCTTGGAACC 2520
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 GACCTCGAGCCGTACTGGGAACGCTGTGCGCGGAGCTGTACCTGCCATTTGAGTCTTCGGAACC 2520

Qy 2521 CCAGGCCCAACCGGTCGCGTCGCCACGCGTCGCCACCCACGGCGGACAGTTGTCCAACCTG 2580
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 CCAGGCCCAACCGGTCGCGTCGCCACGCGTCGCCACGCCCAGGCGGACAGTTGTCCAACCTG 2580

Qy 2581 CGTGCACAGGCACACCGGCCACCCACTGGGCCTTGCTGTCATCGGCTTGCTGATCGGCTTCATCGATGCC 2640
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 CGTGTTACAGGCACACCGGCCACCCACTGGGCCTTGCTGTCATCGGCTTGCTGATCGGCTTCATCGATGCC 2640

Qy 2641 GCAGCCGGTTAATGAGATGCTGGAAGAGCGGACGCCCAACCAAGGTCACCCCCATCCTCCAAGGTTGTT 2700
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 GCAGCCGGTTAATGAGATGCTGGACGCTGGAAGAGCGGACGCCCAACCAAGGTCACCCCCATCCTCCAAGGTTGTT 2700

Qy 2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTGCGGGTGTAGATCCAGCAGACTTTGCTGCGAGAC 2760
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTGCGGGTGTAGATCCAGCAGACTTTGCTGCGAGAC 2760
```

PD   06-JUL-2000.
XX
PF   23-DEC-1998;   98WO-US27301.
XX
PR   23-DEC-1998;   98WO-US27301.
XX
PA   (SINS/) SINSKEY A J.
PA   (LESS/) LESSARD P A.
PA   (WILL/) WILLIS L B.
XX
PI   Sinskey AJ, Lessard PA, Willis LB;
XX
DR   WPI; 2000-465746/40.
DR   P-PSDB; AAB01436.
XX
PT   Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT   carboxylase useful for industrial fermentation processes comprises a
PT   specific nucleotide sequence
XX
PS   Claim 3; Fig 1; 51pp; English.
XX
CC   The pyruvate carboxylase of Corynebacterium glutamicum can be used
CC   for producing amino acids, preferably lysine and glutamic acid in
CC   industrial fermentations and for replenishing oxaloacetate consumed
CC   for biosynthesis during growth. By incorporating the pyruvate
CC   carboxylase gene in expression vectors levels of expression can be
CC   2 - 20 fold higher than in Corynebacterium glutamicum.
XX
SQ   Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 other;

Query Match   97.8%;   Score 3398.8;   DB 21;   Length 3621;
Best Local Similarity   98.6%;   Pred. No. 0;
Matches   3427;   Conservative   0;   Mismatches   47;   Indels   0;   Gaps   0;

```
Qy      1 GTGACTGCTATCACCTTGGCGTCTCTTGTGTTGAAAGGAATAATTACTCTCAGTGTCGACT 60
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    148 GTGACTGCTATCACCCCTTGGCGTCTCTTGTGTTGAAAGGAATAATTACTCTCAGTGTCGACT 207

Qy     61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267

Qy    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327

Qy    181 CGTGAAGAGTCGGGGATCATTCCACCGCTCTTTTGCTTCGAAGCTGTCCGCCATTGGTACT 240
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    328 CGTGAAGAGATCGGGGATCATTCCACCGCTCTTTTGCTTCGAAGCTGTCCGCCATTGGTACC 387

Qy    241 GAAGGCTCACCAGTCAAGGCGTACCTGACATCGACATGAAATTATCGGTGCAGCTAAAAAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 GAAGGCTCACCAGTCAAGGCGTACCTGACATCGACATGAAATTATCGGTGCAGCTAAAAAA 447

Qy    301 GTTAAAGCAGATGCTATTACCCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
          |||||||||||||| |||||||||||||| ||||||||||||||||||||||||||||
Db    448 GTTAAAGCAGATGCCATTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507

Qy    361 CGGAGTGGCGGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGTTCTTGATCTC 420
          ||||||| |||||||||||||||||||||||||||||||||||||| |||||||||||
Db    508 CGGAGTGTGCGGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567

Qy    421 ACCGGTGATAAGTCTCGTGCGGTAACCGCGCCGAACGCGCCGAAGAAGGCTGGTCTGCCG 480
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    568 ACCGGTGATAAGTCTCGTGCGGTAACCGCGCCGAAGAAGGCTGGTCTGCCG 627

Qy    481 GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGCCGTTGAAGGCCAGACTTAC 540
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    628 GAATCCACCCGAGCAAAAACATCGATGAGATCGTTAAAAGCCGTTGAAGGCCAGACTTAC 687

Qy    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGCGGACGCGCGGTATGCGCGCTTTGTTCTTCCA 600
          ||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||||||
Db    688 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGCGGACGCGCGGTATGCGCGGTATTGTTGTTGCTTCA 747

Qy    601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC 660
          ||||||||||| ||||||| ||||||||||||||||||||||||||||||||| |||||||
Db    748 CCTGATGAGCTTCGCAAATTGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807

Qy    661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720
          || ||| |||||||||||||||||||||||||||||||| |||| ||||||||||||||
Db    808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867

Qy    721 CTTGGCGATCGCACTGGAGAAGTGTACACCTTTATGAACGTGACTGCTCACTGCGAGCGT 780
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868 CTTGGCGATCACACTGGAGAAGTGTACACCTTTATGAACGTGACTGCTCACTGCGAGCGT 927

Qy    781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCCTGAT 987

Qy    841 CGCATTTGTCGCGATGCAGTCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    988 CGCATTTGTCGCGATGCAGTCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGGCGCGGGAACC 1047

Qy    901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 960
          |||||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||
Db   1048 GTGGAATTCTTGGTCGATGAAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC 1107

Qy    961 CAGGTTGAGCACACCGTGACTGGAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 CAGGTTGAGCACACCGTGACTGGAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATG 1167

Qy   1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCGACCCAAGATAAGATCAAGACC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCGACCCAAGATAAGATCAAGACC 1227

Qy   1081 CACGTGCAGCACTGCAGTGCCGCCGATCACACCGGAGATCAAACACACGGCTTCCGCCCCA 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 CACGTGCAGCACTGCAGTGCCGCCGATCACACCGGAGATCAAACACACGGCTTCCGCCCCA 1287

Qy   1141 GATACCGGAACTATCACCGCTCACCGGAGTGGCGTTCGTCTTGACGGGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 GATACCGGAACTATCACCGCTCACCGGAGTGGCGTTCGTCTTGACGGGT 1347

Qy   1201 GCAGCTCAGCTGGTCGGTGGCGCAAATCACCCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 GCAGCTCAGCTCAGCTGGTCGGTGGCGCAAATCACCCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407

Qy   1261 TGCCGGTGGTTCCGACTTTGAAACTGCTGTTGTCGTGCAGCGCGCGCGTTGGCTGAGTTC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 TGCCGGTGGTTCCGACTTTGAAACTGCTGTTGTCGTGCAGCGCGCGCGTTGGCTGAGTTC 1467

Qy   1321 ACCGGTCTCGTGTTGCAACCAACATTGGTTTCTGCGGTTCGTGCGGGAAGAGGAC 1380
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1468 ACCGGTGTCTCGGTGTTGCAACCAACATTGGTTTCTGCGGTTCGTGCGGGAAGAGGAC 1527

Qy   1381 TTCACTTCCAAGCGCATGCCACCGGATTATCGGCAGATGTCACCGTGAAC 1440
          |||||| |||||||||||||||||||| |||||| ||||||||||||||||
Db   1528 TTCACTTCCAAGCGCATGCCACCGGATTCATTGCCGATTACTTGGCAGATGTCACCGTGAAC 1587

Qy   1441 CCACCTGCGGATGATGAGCGGGACGCATCCTGGATTACTTGGCAGATGTGCCAACATC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1588 CCACCTGCTGCGGATGATGAGCGGGACGCATCCTGGATTACTTGGCAGATGTCCTAACATC 1647

Qy   1501 AAGCCTCATGTGTCGTCCAAAGGATGTTGCAGCCTCCTATCGATAAGCTGCCAACATC 1560
          ||||||||||||||||||||||| ||||||||||||||||||||||| |||||| |||||
Db   1648 AAGCCTCATGTGTCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707

Qy   1561 AAGGATCTGCCACTGCCGGTTCCCGTGACCGCGGCTTGGCCGCCCAGCCCG 1620
          ||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1708 AAGGATCTGCCACTGCCGGTTCCCGTGACCGCGGCACGGCTTGGCCGCCCAGCCCG 1767

Qy   1621 TTTGCTTCGTGATCTCCGTGAGCAGGACGGCTTTTGTTCTTCCA 1680
          |||||||||||||||||||||||||||||||||||||||||||| |
Db   1768 TTTGCTTCGTGATCTCCGTGAGCAGGACACTGGCTTTTGTTGTTGCTTCA 1827
```

```
Qy   1681 GCACACCAGTCTTTGCTTGGCGACCCGAGTCCGCTCATTCGGCACTGAAGCCTGCGGCGAGAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 GCACACCAGTCTTTGCTTGGCGACCCGAGTCCGCTCATTCGGCACTGAAGCCTGCGGCGAGAG 1887

Qy   1741 GCCGTGCAAAGTCGACTCCTGAGCTTTGTCCGTGGAGGCTTGGGCGGCGGCGGCGACCTAC 1800
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1888 GCCGTGCAAAGTCGACTGACTCCTGAGCTTTGTCCGTGGAGGCTTGGGCGGCGGCGGCGACCTAC 1947

Qy   1801 GATGTGGCGATGCGTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCCGGAG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1948 GATGTGGCGATGCGTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCCGGAG 2007

Qy   1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGATACACC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2067

Qy   1921 CCGTACCCAGACTCCGTTGTTAAGGAAGCTGCCAGCTCCGGCGTCCGGCGTGGAC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2068 CCGTACCCAGACTCCGTTGTTAAGGAAGCTGCCAGCTCCGGCGTCCGGCGTGGAC 2127

Qy   1981 ATCTTCCGCATCTTCGACGCGCCTAACGACGTCTCCCAGATCGTTCAGCAATCGACGCGA 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2128 ATCTTCCGCATCTTCGACGCGCCTAACGACGTCTCCCAGATCGTTCAGCAATCGACGCGA 2187

Qy   2041 GTCCTGGAGACCAACACCCGCGTAGCCGAGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2188 GTCCTGGAGACCAACACCCGCGTAGCCGAGTGGCTATGGCTTATTCTGGTGATCTCTCT 2247

Qy   2101 GATCCAAATGAAAAGTCTACACCCGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2248 GATCCAAATGAAAAGTCTACACCCGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2307

Qy   2161 AAGTCTGGGCGCTCACATTCTGGCCATTCGGTGTCTGCTTCGCCCAGCTGCG 2220
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2308 AAGTCTGGGCGCTCACATCTCTGGCCATTCGGTGTCTGCTTCGCCCAGCTGCG 2367

Qy   2221 GTAACCAAGCTGGTCACCGGCCACTGCGCCGTGAATTCGATCGCACGTGCACACC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2368 GTAACCAAGCTGGTCACCGGCCACTGCGCCGTGAATTCGATCGCACGTGCACACC 2427

Qy   2281 CACGACACTGCGGGGTGGCCAGTTGGCTGCAACTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2428 CACGACACTGCGGGGTGGCCAGTTGGCTGCAACTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487

Qy   2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGACCACCTCCGCCATCCGTCTGCGCC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGACCACCTCCGCCATCCGTCTGCGCC 2547

Qy   2401 ATTGTTGCTGCATTCGGCGACACACCGGTTGAGCCTCGGAGGCTGTTTCT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2548 ATTGTTGCTGCATTCGGCGACACACCGGTTGAGCCTCGGAGGCTGTTTCT 2607

Qy   2461 GACCTCGGAGCTGGGAAGCTGTGCGCGGACGTGCTGGAAGTCTCCGAACC 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2608 GACCTCGGAGCTGGGAAGCTGTGCGCGGACTGCTGGAAGTCTCCGAACC 2667

Qy   2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGCGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGCGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727

Qy   2581 CGTGCACAGGCCACCGCACTGGGCCTTGATGCGTTCGAGCTTCATCGAAGACAACTAC 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2728 CGTGCACAGGCCACCGCACTGGGCCTTGATGCGTTCGAACTCATCGAAGACAACTAC 2787

Qy   2641 GCAGCCGTTAATGAGATGCTGGACGCGCCCAACCAAGGTCACCCATCCTCCAAGGTTGTT 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2788 GCAGCCGTTAATGAGATGCTGGGACGCGCCCAACCAAGGTCACCCATCCTCCAAGGTTGTT 2847

Qy   2701 GGGACCTCGCACTCACCTGGTTGCTGCGGGTGAGATCGACAGCTTGCTCAGAC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2848 GGGACCTCGCACTCACCTCCGGTTGCTGCGGGTGCGGGTGAGATCGACAGCTTGCTCAGAC 2907

Qy   2761 CCACAAAGTACGACGACATCCGTGCTCGTCATCGCGGTTCCTGCGGCGGCGGCGCTTGGTAAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2908 CCACAAAGTACGACGACATCCGTGCTCGTCATCGCGGTTCCTGCGGCGGCGGCGCTTGGTAAC 2967

Qy   2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCACTGGAAGGCCGTTCCGAAGGC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCACTGGAAGGCCGTTCCGAAGGC 3027

Qy   2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCACCTCGACGCTGCACCTCGATGATGATTCC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCACCTCGACGCTGCACCTCGATGATGATTCC 3087

Qy   2941 AAGGAACGTCGCAACAGCCTCAACCGCTGCGTTCCCGAAGCAACCGAAGAGTTCCTC 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3088 AAGGAACGTCGCAATAGCCTCAACCGCTGCGTTCCCGAAGCAACCGAAGAGTTCCTC 3147

Qy   3001 GAGCACCGTCGCCGCTTCGGCAACTCACCTCGGCGTGATGATCGTGATGATTCTTCTACGGA 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3148 GAGCACCGTCGCCGCTTCGGCAACTCACCTCGGCGTGATGATCGTGATGATTCTTCTACGGC 3207

Qy   3061 CTGGTCGAGGGCCGGCCGCCGGAGACTTTGATCGCGTCCAGATGTGCGCACCCCACTGCTTGTT 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 CTGGTCGAGGGCCGGCCGCCGGAGACTTTGATCGCGTCCAGATGTGCGCACCCCACTGCTTGTT 3267

Qy   3121 CGCCTGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGCGCCAACGTC 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 CGCCTGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGCGCCAACGTC 3327

Qy   3181 AACGGCCAGATCCGCCCAATGCGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCCAACC 3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   3328 AACGGCCAGATCCGCCCAATGCGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCCAACC 3387

Qy   3241 GCAGAAAAGGCAGATTCCTCCAACAGAGGGCCATGGTGCTGCACCATTCGCTGTGTTGTC 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 GCAGAAAAGGCAGATTCCTCCAACAGAGGGCCATGGTGCTGCACCATTCGCTGTGGTGTC 3447

Qy   3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTGCAATCATCGAG 3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTGCAATCATCGAG 3507

Qy   3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGACGGCAAGCGCGGTGTG 3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATGCGCGGTTGTG 3567

Qy   3421 GTTCCTGCTGCAACGAAGGTGGAAGTGCGACTTGATCGTCGTGTTTCCTAA 3474
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 GTTCCTGCTGCAACGAAGGTGGAAGTGCGACTTGATCGTCGTGTTTCCTAA 3621
```

PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI; 2001-122330/13.
DR P-PSDB; AAB67129.
XX
PT Novel nucleic acid encoding pyruvate carboxylase from Corynebacterium
PT glutamicum, for replenishing oxaloacetate consumed during lysine and
PT glutamic acid production in industrial fermentations -
XX
PS Claim 2; Column 23-30; 29pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC Corynebacterium glutamicum pyruvate carboxylase protein. This is an
CC enzyme in the anaplerotic pathway. It can be used in the replenishment of
CC oxaloacetate consumed during lysine and glutamic acid production in
CC industrial fermentation.
XX
SQ Sequence 3621 BP; 782 A; 1037 C; 1002 G; 800 T; 0 other;

Query Match   97.8%; Score 3398.8; DB 22; Length 3621;
Best Local Similarity   98.6%; Pred. No. 0;
Matches 3427; Conservative   0; Mismatches   47; Indels   0; Gaps   0;

Qy   1 GTGACTGCTATCACCCTTGGCGGTCTCCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148 GTGACTGCTATCACCCTTGGCGGTCTCCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207

Qy   61 CACCATCTTCAACGCTTCAAGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 CACCATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267

Qy   121 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 GCGGTCCGTGCTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC 327

Qy   181 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387

Qy   241 GAAGGCTCACCAGTCAAGGCGTACTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GAAGGCTCACCAGTCAAGGCGTACTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447

Qy   301 GTTAAAGCAGAGATGCTATTTACCCGGGATATGTCCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 GTTAAAGCAGAGATGCCATTTACCCGGGATATGTCCTTCCTGTCTGAAAATGCCCAGCTTGCC 507

Qy   361 CGGCAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 CGGCAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC 567

Qy   421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAGAAGGCGTGGTCTGCAGTTTTGGCG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAGAAGGCTGGTCTGCAGTTTTGGCG 627

Qy   481 GAATCCACCCCGAGCAAAAACAATCGATGACATCGTTAAAAGCGTGAAGGCCAGACTTAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GAATCCACCCCGAGCAAAAACAATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC 687

Qy   541 CCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGGCGCTTGTTTCTTCA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 CCATTCTTTGTGAAGGCAGTTGCCGGTGGTGGCGGACGCGGTATGGCGTTTGTTGCTTCA 747

Qy   601 CCTGATGAGCTCCGCAAATTGCCAACAGAAGCATCCGTGAAGCTGAACGGCCATTCGGC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 CCTGATGAGCTTCGCAAATTAGCAACAAGAAGCATCTCGTGAAGCTGAACGGCCGC 807

Qy   661 GACGGTTCGGTATATGTCGAACGTGCTGTAACCCCCAGCACATTGAAGTGCAGATC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 GATGGCGCGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867

Qy   721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
            ||||||||||||||||||||||||||||||||||||||||

Db   868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTGCTCACTGCAGCGT 927

Qy   781 CGTCACCAAAAAGTGTCGAAATTGGGCCAGCACAGCCATTGGATCTCCAGAACTGCGTGAT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 CGTCACCAAAAAGTGTCGAAATTGGGCCAGCACAGCCATTGGATCTCCAGAACTGCGTGAT 987

Qy   841 CGCATTTGTGCGGATGCAGTAAAGTTCTCGCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 CGCATTTGTGCGGATGCAGTAAAGTTCTCGCGCTCCATTGGTTACCAGGGCGCGGGAAACC 1047

Qy   901 GTGGAATTCTTGGTGCGATGAAAAGGCCAACCACGTTTCATCGAAATGACCACGTATC 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048 GTGGAATTCTTGGTGCGATGAAAAGGGCAACCACCGTCTTCATCGAAATGACCACGTATC 1107

Qy   961 CAGGTTGAGCACACCGTACTGAAGAGTCACCGAGTGGACCTGTGAAGGCGCAGATG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 CAGGTTGAGCACACCGTGACTGAAGAGTCACCGAGTGGACCTGTGAAGGCGCAGATG 1167

Qy   1021 CGCTTGGCTGCTGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 CGCTTGGCTGCTGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227

Qy   1081 CACGGTGCAGCACTGCAGTGCGCGCATCACCGGAGAATCCAAACAACGGCTTCCGCCCA 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 CACGGTGCAGCACTGCAGTGCGCCGCATCACCACCACGGAGAATCCAAACAACACGGCTTCCGCCCA 1287

Qy   1141 GATACCGGAACTATCACCGCGTACCGTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 GATACCGGAACTATCACCGCGTACCGTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT 1347

Qy   1201 GCAGCTCAGCTCCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 GCAGCTCAGCTCCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC 1407

Qy   1261 TGCCGGTGTTCCGACTTTGAAACTGCTGTTGCTGTCGTGCACAGCGCCGGTTGGCTGAGTTC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 TGCCGGTGTTCCGACTTTGAAACTGCTGTTGCTGTCGTGCACAGCGCCGGTTGGCTGAGTTC 1467

Qy   1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCCGGATCACCCCACCTCCTTCAGGCT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCCGGATCACCACCGCCACCTCCTTCAGGCT 1527

Qy   1381 TTCACTTCCAAGGCATCGCACCGGATTTATCGGGCAGCAATTGGATCACGTGAAC 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528 TTCACTTCCAAGGCATCGCACCGGATTTATCGGGCAGCAATTCATTGCGGATGTTCACCGTGAAC 1587

Qy   1441 CCACCTGCGATGATGAGGACGCATCGATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1588 CCACCTGCTGATGATGAGGACGCATCGATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1647

Qy   1501 AAGGCTCATGGTGTGCGTCCAAAGGATGTTGAGCACCAATCGATAAGCTGCCCAACATC 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648 AAGGATCTGCCACTGCCTGTGCGTCCCACGCGGTTCCCGTGACCGGCCTGAGCAGCTGCCTAACATC 1707

Qy   1561 AAGGAATCTGCCACTGCCTGTCCCGTGACCGGCTTGGCCTGAAGCGCCAGCCGCCG 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708 AAGGATCTGCCACTGCCTGTCCCGTGACCGGCTTGGCCTGAAGCAGCTGGCCAGCCGCCG 1767

Qy   1621 TTTGCTCGTGATCTCCGTGACGGTGGTGGCGAGAGGCGACGGCCGTTACTGATACCGCGAT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1768 TTTGCTCGTGATCTCCGTGACGGTGGTGGCGAGCAGGACGGCCGTTACTGATATACCGCGAT 1827

Qy   1681 GCAACCAGTCTTTGCTTTGCTTCGTGAAGCTGAACGGCCATTCGGC 1740
            |||||||||||||||||||||||||||||||||||||||||||||
Db   1828 GCACACCAGTCTTTGCTTTGCTTCGTGAAGCTGAACGGCCTGCGAGAG 1887

Qy   1741 GCCGTGCAAAGCTGACTTCCTGAGCTTTTGTCCGTGGAGGCCTGGGCGCGCGACTTAC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1888 GCCGTGCAAAGCTGACTTCCTGAGCTTTTGTCCGTGGAGGCCTGGGCGCGGCGCGACTTAC 1947

Qy   1801 GATGTGGCGATGCGTTCCTTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCCGCGAG 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1948 GATGTGGCGATGCGTTCCTTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCCGCGAG 2007

Left column:

```
Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCGCAACACCGTGGGATACACC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCGCAACACCGTGGGATACACC 2067

Qy  1921 CCGTACCCAGACTCCGTCTGCGCGCGGTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC   1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2068 CCGTACCCAGACTCCGTCCTGCGCGCGGTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGAC  2127

Qy  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA  2187

Qy  2041 GTCCTGGAGACCAACACCGGGTAGCCGAGGTGGCTATGGCTTATTCTGTGATCTCTCT    2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2188 GTCCTGGAGACCAACACCGGGTAGCCGAGGTGGCTATGGCTTATTCTGTGATCTCTCT   2247

Qy  2101 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGCAGAGGAGATCGTC    2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2248 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGCAGAGGAGATCGTC   2307

Qy  2161 AAGTCTGGCGCTCACATTCTGGCCATTAAGGATATGGCTCGCTTCGCTCGCCCAGCTGCG  2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2308 AAGTCTGGCGCTCACATCTGGCCATTAAGGATATGGCTCGCTTCGCTCGCCCAGCTGCG   2367

Qy  2221 GTAACCAAGCTGGTCACCGACCGGCCGTGAATTCGATCCAGCTGCACGTGCACAC       2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2368 GTAACCAAGCTGGTCACCGACCGGCCGTGAATTCGATCCAGCTGCACGTGCACACC      2427

Qy  2281 CACGACACTGCGGGGTGCCAGTTGGCTACCTACTTTGCTCGAGCTGGTGCAGAT       2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2428 CACGACACTGCGGGGTGCCAGTTGGCTACCTACTTTGCTCGAGCTGGTGCAGAT      2487

Qy  2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCAGCATCCCTGTCTGCC    2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCAGCATCCCTGTCTGCC   2547

Qy  2401 ATTGTTGCTGCTGCATTCGCGCACCACCCGGTTTGAGCCTCGGCGCTGTTTCT       2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2548 ATTGTTGCTGCTGCATTCGCGCACCACCCGGTTTGAGCCTCGGCGCTGTTTCT      2607

Qy  2461 GACCTCGAGCTACTGGGAAGTGCGCGCGGACTGTACCTGACCTGAGTCTCGAACC    2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2608 GACCTCGAGCTACTGGGAAGTGCGCGCGGACTGTACCTGACCTGAGTCTCGAACC   2667

Qy  2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACCTG  2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2668 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACCTG  2727

Qy  2581 CGTGCACAGGCCACCCACTGGGCCTTGCTCGTGATCGCTTCGAGCGTCATCGAAGACAACTAC  2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2728 CGTGCACAGGCCACCCACTGGGCCTTGCTCGTGATCGCTTCGAGCGTCATCGAAGACAACTAC  2787

Qy  2641 GCAGCCGTTAATGAGATGCGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2788 GCAGCCGTTAATGAGATGCGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2847

Qy  2701 GGGGACCTCGCACTCCACTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCAGAC    2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2848 GGGGACCTCGCACTCCACTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGAT   2907

Qy  2761 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGCGGAGCTTGGTAAC   2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2908 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCGCGCGGAGCTTGGTAAC  2967

Qy  2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCCCGGCGACTGGAAGGCCGCTCCGAAGGC  2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCCGGCCGCTCCCGAAGGC          3027

Qy  2881 AAGGCACCCTGACGGAAGTTCCTGAGGAGCAGGCGCACCTCGACGCTGATGATTCC       2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3028 AAGGCACCCTGACGGAAGTTCCTGAGGAGCAGGCGCGCACCTCGACGCTGATGATTCC   3087
```

Right column:

```
Qy  2941 AAGGAACGTGCAACAGCCTCAACCGCCTGCTGTTCCCGAGCCAACCGAAGAGTTCCTC  3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3088 AAGGAACGTGCAATAGCCTCAACCGCCTGCTGTTCCCGAGCCAACCGAAGAGTTCCTC 3147

Qy  3001 GAGCACCGTCGCCGCTTCGGCAACTCCTGCTGCTTCGCGTGATGATCGTGAATTCTTCTACGGA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3148 GAGCACCGTCGCCGCTTCGGCAACTCCTGCTGCTTCGCGTGATGATCGTGAATTCTTCTACGGC 3207

Qy  3061 CTGGTCGAGGGCGGCGGAGACTTGATCGCGCCTGCCCAGATGTGCGCACCCACTGCTTGTT  3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3208 CTGGTCGAAGGCGGCGGAGACTTTGATCGCGCCTGCCCAGATGTGCGCACCCACTGCTTGTT 3267

Qy  3121 CGCCTGGATGCGATCTCTGAGCCAGCAGCGATAAGGGTATGCCGCAATGTTGTGGCCAACGTC  3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3268 CGCCTGGATGCGATCTCTGAGCCAGCAGCGATAAGGGTATGCCGCAATGTTGTGGCCAACGTC 3327

Qy  3181 AACGGCACAGATTCCGCCCAATGCGTGCGTGACGTCCGTTGAGTCGTCGTCACCGCAACC  3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3328 AACGGCCAGATTCCGCCCAATGCGTGCGTGACGTCCGTTGAGTCGTCGTCACCGCAACC   3387

Qy  3241 GCAGAAAAGGCAGATTCCTCCAACAAGGCCATGTTCTGCCACCATTCGCTCGTGTGTC   3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTCTGCCACCATTCGCTCGTGTGTC  3447

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGGAGATGCAGTGCGCAATCATCGAG 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTGAGGTCAAGGCTGGGAGATGCAGTCAGTCGCAATCATCGAG 3507

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGACGGCAAGATTGAACGCGTTGTG  3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3508 GCTATGAAGATGGAAGCAACAATCACTGCTTCTCGTTGTGACGGCAAAATCGATCGCGGTTGTG 3567

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGCGACTTGATCGTCGTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGGTCGTCGTTCCTAA 3621
```

RESULT 4

AAH68526
ID   AAH68526 standard; DNA; 349980 BP.
XX
AC   AAH68526;
XX
DT   26-SEP-2001 (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 7061.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI   Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR   WPI; 2001-376931/40.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -

```
XX
PS   Disclosure; SEQ ID NO: 7061; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a nucleic acid described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence 349980 BP; 79274 A; 90638 C; 98727 G; 81341 T; 0 other;

Query Match              97.8%;  Score 3398.8;  DB 22;  Length 349980;
Best Local Similarity    98.6%;  Pred. No. 0;
Matches 3427;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy      1 GTGACTGCTATCACCCTTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTGACT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105160 GTGACTGCTATCACCCTTGGCGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTGACT 105219

Qy     61 CACACATCTTCAACGCTTCAGCATTCAAAAAGATCTTGGTAGCAAAACGCGGCGAAATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105220 CACACATCTTCAACGCTTCAGCATTCAAAAAGATCTTGGTAGCAAAACGCGGCGAAATC 105279

Qy    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105280 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCC 105339

Qy    181 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db 105340 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 105399

Qy    241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105400 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAA 105459

Qy    301 GTTAAAGCAGATGCTATTACCCGGGATAACGGCTTTTATTGGCCCAACCCCAGCTTGCC 360
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 105460 GTTAAAGCAGATGCATTTACCCGGGATAACGGCTTTTATTGGCCCAACCCCAGCTTGCC 105519

Qy    361 CGCGAGTGCGCGGAAAACGCCATTACTTTATTGGCCCAGAGGTTCTTGATCTC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105520 CGCGAGTGCGCGGAAAACGCCATTACTTTATTGGCCCAGAGGTTCTTGATCTC 105579

Qy    421 ACCGGTGATAAGTCGTCGGTAACCGCCGAGAAGGCTGGTCTGCCAGTTTTGGCG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105580 ACCGGTGATAAGTCGTCGGTAACCGCCGAGAAGGCTGGTCTGCCAGTTTTGGCG 105639

Qy    481 GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGGCGCTGAAGGCCAGACTTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105640 GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGGCGCTGAAGGCCAGACTTAC 105699

Qy    541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCGACAGAGAAGCGTCGTATGCGCGTTCTTCA 600
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 105700 CCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGCGACAGAGAAGCGTCGTATGCGCGTTTGTGCTTCA 105759

Qy    601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGGCATTCGGC 660
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||
Db 105760 CCTGATGAGCTCCGCAAATTCGCAACAGAAGCATCTCGTGAAGCTGAAGCGGGCTTTCGGC 105819

Qy    661 GACGGTTCGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720
          |||||||||||||||||||||||||||||||||||||| |||||||| |||||||||||
Db 105820 GACGGTTCGTATATGTCGAACGTGCTGTGATTAACCCTCAGCACATATTGAAGTGCAGATC 105879

Qy    721 CTTGGCGATCGCACTGAGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
          |||| |||
Db 105880 CTTGGCGATCACACTGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 105939

Qy    781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCTGAT 840
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db 105940 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCTGAT 105999

Qy    841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db 106000 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCCGCTCCATTGGTTACCAGGGGCGCGGGAACC 106059

Qy    901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 960
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db 106060 GTGGAATTCTTGGTCGATGATGAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 106119

Qy    961 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGTGAAGGCGCAGATG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106120 CAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGTGAAGGCGCAGATG 106179

Qy   1021 CGCTTGGCTGCTGGTGCAACCTTGACGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
          |||||| || ||||||||||||||||||||||||||||| |||||||||||||||||||
Db 106180 CGCTTGGCTGCTGGTGCAACCTTGACGAATTGGGTCTGACCCAAGATAAGATCAAGACC 106239

Qy   1081 CACGGTGCAGCAGACGCACTGCAGTCGCGTCACCGCTCACCGGAGATCAAAACAACGGCTTCCGCCCA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106240 CACGGTGCAGCAGACGCACTGCAGTCGCGTCACCGCTCACCGGAGATCAAAACAACGGCTTCCGCCCA 106299

Qy   1141 GATACCGGAACTATCACCGCGTACCGCTCACCGGAGTCGGCGTTCGTCTTGACGGT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106300 GATACCGGAACTATCACCGCGTACCGCTCACCGGAGTCGGCGTTCGTCTTGACGGT 106359

Qy   1201 GCAGCTCAGCTGCGTCGGTGGCGAAATCACCGCCACACTTTGACTCCATGCTGGTGAAAATGACC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106360 GCAGCTCAGCTGCGTCGGTGGCGAAATCACCGCCACACTTTGACTCCATGCTGGTGAAAATGACC 106419

Qy   1261 TGCCGTGGTTCCGACTTTGAAACTGTCGTTGCTGCGCACGCGCGCGGTTGGCTGAGTTC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106420 TGCCGTGGTTCCGACTTTGAAACTGTCGTTGCTGCGCACGCGCGCGGTTGGCTGAGTTC 106479

Qy   1321 ACCGGTGCTGGTGTTCGAACCAACCATTGGTTCTTGCGTGCGTCGATCACCGCGATCACCGCGGAGGAC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106480 ACCGGTGCTGGTGTTCGAACCAACCATTGGTTTTGCGTGCGTCGATCACCGCGATCACCGCGGAGGAC 106539

Qy   1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATTACTTGCCAGATGTCACCGTGAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106540 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATTACTTGCCAGATGTCACCGTGAAC 106599

Qy   1441 CCAACCTGCGATGATGAGCAGGAGGACGCATCCGAGATGGTGCAGCACCAATCGATAAGCTGCCAACATC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106600 CCACCTGCGATGATGAGCAGGAGGACGCATCCGAGATGGTGCAGCACCAATCGATAAGCTGCCTAACATC 106659

Qy   1501 AAGGCCTATAAGTCCGTCCCGAGAAGAGGCTGGTCTGCCAGTTTTGGCG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106660 AAGGCCTTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 106719

Qy   1561 AAGGATCTGCCACTGCCCACGCGGTTCCCGTGACCGCGGTCTGCCAGCCGCGCG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106720 AAGGATCTGCCACTGCCCACGCGGTTCCCGTGACCGCGGTCTGCCAGCCGCGCG 106779

Qy   1621 TTTGCTGTGATCTCCGTGCTGCGACGAGGACCACTGGCAGTTACTGATACCACCTTCCGCGAT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106780 TTTGCTGTGATCTCCGTGCTGCGACGAGGACCACTGGCAGTTACTGATACCACCTTCCGCGAT 106839

Qy   1681 GCACACCAGTCTTTGCTGCGGCAACAGAAGCATCTCGTGAAGCGCGGGCATTCGGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106840 GCACACCAGTCTTTGCTGCTTCTGCGGCACCCGAGCTTGGGCTCGGCGGCGACCTAC 106899

Qy   1741 GCCGTGCGCAAAGCTGACTCCGAGCTTTGTTGCCGTGGACGCTTTTGTCGTGGGGCGGCGACCTAC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106900 GCCGTGCGCAAAGCTGACTCCGAGCTTTGTCGTGGACGCTTTTGTCGTGGGGGCGGCGACCTAC 106959

Qy   1801 GATGTGGCGATGCCGTTTCCTCTTTGAGGATCCGTGGGACACAGGCTCGACGAGCTGCGCGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 106960 GATGTGGCGATGCCGTTTCCTCTTTGAGGATCCGTGGGACACAGGCTCGACGAGCTGCGCGAG 107019
```

```
Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTCGCGGGCCGCAACACCGTGGGATACACC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107020 GCGATGCCGAATGTAAACATTCAGATGCTGCTCGCGGGCCGCAACACCGTGGGATACACC 107079

Qy  1921 CCGTACCCAGACTCCGTCGCGCCGGCGTTTGTTAAGGAAGCTGCCAGTCCGGCGTGGAC 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107080 CCGTACCCAGACTCCGTCGCGCCGGCGTTTGTTAAGGAAGCTGCCAGTCCGGCGTGGAC 107139

Qy  1981 ATCTTCCGCATCTTCACGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107140 ATCTTCCGCATCTTCACGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 107199

Qy  2041 GTCCTGGAGACCAACACCGGGTAGCCGAGGTGGCTATGCCGAGGTGGCTATTCTGGTGATCTCTCT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107200 GTCCTGGAGACCAACACCGGGTAGCCGAGGTGGCTATGCCGAGGTGGCTATTCTGGTGATCTCTCT 107259

Qy  2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107260 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 107319

Qy  2161 AAGTCTGGCCGCTCACATTCTGCCATTAAGGATATGGCTGGTCTGCTTCGCGCCAGCTGCG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107320 AAGTCTGGCCGCTCACATTCTGCCATTAAGGATATGGCTGGTCTGCTTCGCGCCAGCTGCG 107379

Qy  2221 GTAACCAAGCTGGTCACGGACTGGTCACCGCCACTGCGCCGTGAATTCGATCTGCACGTGCACACC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107380 GTAACCAAGCTGGTCACGGACTGGTCACCGCCACTGCGCCGTGAATTCGATCTGCACGTGCACACC 107439

Qy  2281 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTTGCTCGCGAGCTCAAGCTGGTGCGAGAT 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107440 CACGACACTGCGGGGTGGCCAGTTGGCTACCTACTTTGCTCGCGAGCTCAAGCTGGTGCGAGAT 107499

Qy  2341 GCTGTTGACGGTGCTTCCGGACCACTGTCTCTGGCCACCACTGTCTCTGGCCAGCCATCCCGTCTGCC 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107500 GCTGTTGACGGTGCTTCCGGACCACTGTCTCTGGCCACCACTGTCTCTGGCCAGCCATCCCGTCTGCC 107559

Qy  2401 ATTGTTGCTGCATTCGCGCACACCGTCGCGCCAACACCGGTTTGAGCCTTCGAGGCTTGTTCT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107560 ATTGTTGCTGCATTCGCGCACACCGTCGCGCCAACACCGGTTTGAGCCTTCGAGGCTTGTTTCT 107619

Qy  2461 GACCTCGAGCGCGTACTGGGAAGCGTGTGCGCGGACTGTACCTGCCATTGAGTCTCGAAAC 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107620 GACCTCGAGCGCGTACTGGGAAGCGTGTGCGCGGACTGTACCTGCCATTGAGTCTCGAAAC 107679

Qy  2521 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107680 CCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 107739

Qy  2581 CGTGCACAGGCCACCGCACTGGCCTTGCTGTGATCGCTCATCGAAGCAACAACTAC 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107740 CGTGCACAGGCCACCGCACTGGCCTTGCTGTGATCGCTCATCGAAGCAACAACTAC 107799

Qy  2641 GCAGCCGTTAATGAGATGCGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107800 GCAGCCGTTAATGAGATGCGTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 107859

Qy  2701 GGGGACCTCGCACTCCACCTCGTTGGTGCGGGTGAGATCCAGCAGACTTTGCTGCCGACAC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 107860 GGGGACCTCGCACTCCACCTCGTTGGTGCGGGTGAGATCCAGCAGACTTTGCTGCCGACAC 107919

Qy  2761 CCACAAAGTACGACATCCCAGACTCTGTCATCGCCGAGCTTCGGTAAC 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 107920 CCACAAAGTACGACATCCCAGACTCTGTCATCGCCGAGCTTCGGTAAC 107979

Qy  2821 CCTCCAGGTGCTGGCCAGACACCGGTGCGGAAGGCCGCTCCGAAGGC 2880
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 107980 CCTCCAGGTGCTGGCCAGACACCGGTGCGGAAGGCCGCTCCGAAGGC 108039

Qy  2881 AAGGCACCTCTGACCGGAAGTTCCTGACGGAAGAGCAGGCGCACCTTCGACGCTGATGATTCC 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108040 AAGGCACCTCTGACCGGAAGTTCCTGACGGAAGAGCAGGCGCACCTTCGACGCTGATGATTCC 108099

Qy  2941 AAGGAACGTCGCAACAGCCTCAACCGCTCGTGTTCCGAAGCCAACCGAAGAGTTCCTC 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108100 AAGGAACGTCGCAATAGCCTCACCGCTCGTGTTCCGAAGCCAACCGAAGAGTTCCTC 108159

Qy  3001 GAGCACCGTCGCCGCTTCGGCAACACCTCGGCTGCGGATCGGATCGGAATTCTTCTACGGA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108160 GAGCACCGTCGCCGCTTCGGCAACACCTCGGCTGCGGATCGGATCGGAATTCTTCTACGGC 108219

Qy  3061 CTGGTCGAGGGCGGCAGACTTGATCGCGCCAGATGTGCGCACCCACTGCTTGTT 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108220 CTGGTCGAAGGCCGGCAGACTTTGATCGCGCCAGATGTGCGCACCCACTGCTTGTT 108279

Qy  3121 CGCCTGGATGCGGATCTCTGAGCCAGACGATAAGGGTATGCGGCAATGTTGTGGCCAACGTC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108280 CGCCTGGATGCGGATCTCTGAGCCAGACGATAAGGGTATGCGGCAATGTTGTGGCCAACGTC 108339

Qy  3181 AACGGCCAGATCCGCCCAATGCGTGCGTTGAGTCTCCGTTGAGTCGTGTCCACCGCAACC 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108340 AACGGCCAGATCCGCCCAATGCGTGCGTTGAGTCTCCGTTGAGTCGTGTCCACCGCAACC 108399

Qy  3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACAAGTTGCTGCTGTGTTGTC 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108400 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACAAGTTGCTGCTGTGTTGTC 108459

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108460 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAG 108519

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAGATTGAACGCGGTTGTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108520 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCAAGATTGAACGCGGTTGTG 108579

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTTCCTAA 3474
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 108580 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTTCCTAA 108633
```

RESULT 5
AAX24102
ID   AAX24102 standard; DNA; 3728 BP.
XX
AC   AAX24102;
XX
DT   30-JUN-1999 (first entry)
XX
DE   C. glutamicum pyruvate carboxylase genomic DNA.
XX
KW   Pyruvate carboxylase; amino acid production; lysine production;
KW   threonine production; homoserine production; glutamate production;
KW   arginine production; feed additive; condiment; pharmaceutical;
KW   fine chemical; ss.
XX
OS   Corynebacterium glutamicum.
XX
FH   Key   Location/Qualifiers
FT   CDS   165..3587
FT   /*tag= a
FT   /product= "pyruvate carboxylase"
XX
PN   DE19831609-A1.
XX
PD   15-APR-1999.
XX
PF   14-JUL-1998; 98DE-1031609.
XX
PR   04-OCT-1997; 97DE-1043894.
XX
PA   (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI   Eikmanns B, Peters-Wendisch P, Sahm H;
XX
DR   WPI; 1999-245521/21.
DR   P-PSDB; AAW93971.

```
XX
PT   Increasing microbial production of specific amino acids by
PT   increasing activity or expression of pyruvate carboxylase
XX
PS   Claim 16; Page 8-11; 18pp; German.
XX
CC   This invention describes the isolation of a pyruvate carboxylase
CC   from Corynebacterium glutamicum which is used in a novel method for
CC   production of lysine, threonine, homoserine, glutamate and/or arginine,
CC   variously useful as feed additives, condiments, pharmaceuticals and
CC   intermediates for fine chemicals. Increasing pyruvate carboxylase
CC   activity increases the yield of microbial production of amino acids
CC   of the aspartate and/or glutamate families, e.g. about 50% more lysine,
CC   40% more threonine and 150% more homoserine are secreted into the
CC   culture medium.
XX
SQ   Sequence 3728 BP; 813 A; 1059 C; 1035 G; 821 T; 0 other;

Query Match           97.8%;  Score 3397.2;  DB 20;  Length 3728;
Best Local Similarity 98.6%;  Pred. No. 0;
Matches 3426;  Conservative 0;  Mismatches 48;  Indels 0;  Gaps 0;
```

```
Qy     1 GTGACTGCTATCACCCTTGGCCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT   60
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   114 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT  173

Qy    61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC  120
         ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   174 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC  233

Qy   121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   234 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCC  293

Qy   181 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTCGAAGCTGTCCGCATTGGTACT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   294 CGTGAAGATCGGGGATCATTCACCCGCTCTTTGCTTCTCGAAGCTGTCCGCATTGGTACC  353

Qy   241 GAAGGCTCACCAGTCAAGGCGTACCTGACATCGATGAATTATCGGTGCAGCTAAAAAA   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   354 GAAGGCTCACCAGTCAAGGCGTACCTGACATCGATGAATTATCGGTGCAGCTAAAAAA   413

Qy   301 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC  360
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   414 GTTAAAGCAGATGCTATTTACCCGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCT  473

Qy   361 CGCGAGTGTGCGCGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGACTC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db   474 CGCGAGTGTGCGCGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC  533

Qy   421 ACCGGTGATAAGTCTCGTGCGGTAACCGCGCCGGAGAAGGCTGGTCTGCCAGTTTTGGCG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   534 ACCGGTGATAAGTCTCGTGCGGTAACCGCGCCGGAGAAGGCTGGTCTGCCAGTTTTGGCG  593

Qy   481 GAATCCACCCCAGCAAAAACATCGATGACATCGTAAAGCGCTGAAGGCCAGACTTAC   540
         |||||||||||||||||||||||||||| ||||| |||||||||||||||||||||||
Db   594 GAATCCACCCCAGCAAAAACATCGATGAGATCGTTAAAGCGCTGAAGGCCAGACTTAC   653

Qy   541 CCCATCTTTGTAAAGCCAGTTGTACACCTTTATGAACGTGACACTGCTCACTCCAGCGT  600
         ||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db   654 CCCATCTTTGTGAGGCCAGTTGTACACCTTTATGAACGTGACACTGCTCACTCCAGCGT  713

Qy   601 CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC  660
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||| ||||
Db   714 CCTGATGAGCTCCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCGTTCGGC  773

Qy   661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC  720
         || || |||||||||||||||||||||||||||||||||| |||| |||||||||||||
Db   774 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCCTCAGCATATTGAAGTGCAGATC  833

Qy   721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTCCAGCGT  780
         |||||||||| ||||||||||||||||| |||||||||||||||||||||||||||||||
Db   834 CTTGGCGATCAACACTGGAGAAGTTGTACAACCTTTATGAACGTGACTGCTCACTCCAGCGT  893
```

```
Qy   781 CGTCACCAAAAAGTTGTCGAAATTGCGGCCAGCACAGCATTGGATCCAGAACTGCGTGAT  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   894 CGTCACCAAAAAGTTGTCGAAATTGCGGCCAGCACAGCATTGGATCCAGAACTGCGTGAT  953

Qy   841 CGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCCCGTCCATTGGTTACCAGGGCGCGGGAACC  900
         |||||| ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   954 CGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCCCGTCCATTGGTTACCAGGGCGCGGGAACC  1013

Qy   901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC  960
         |||||||||||||||||| |||||||||||||||||| |||| |||||||||||||||||
Db  1014 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC  1073

Qy   961 CAGGTTGAGCACACCGTGACTCAAGAAGTCACCGAGGTGACCTGGTGAGGCGCAGATG  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1074 CAGGTTGAGCACACCGTGACTCAAGAAGTCACCGAGGTGACCTGGTGAGGCGCAGATG  1133

Qy  1021 CGCTTGGCTGCGTGGTGCAACCTTGAACGAATTGGGTTCTGACCCAAGATAAGATCAAGACC  1080
         |||||||||||||||||||||||||| |||||||||||||||||||| |||||||||||||
Db  1134 CGCTTGGCTGCGTGGTGCAACCTTGAACGAATTGGGTTCTGACCCAAGATAAGATCAAGACC  1193

Qy  1081 CACGGTGCAGCACTGCAGTGCCGTCATCCACCACGGAGATCCAAACAACGGCTTCCGCCCA  1140
         ||||||||||||||||||||| |||||||||||||| |||||||||||||||||||||| |
Db  1194 CACGGTGCAGCACTGCAGTGCCGTCATCCACCACGAGATCCAAACAACGGCTTCCGCCCA  1253

Qy  1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGGCGAGCTGGGCGTTCGTCTTGACGGT  1200
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  1254 GATACCGGAACTATCACCGCGTACCGCTCACCAGGGCGAGCTGGGCGTTCGTCTTGACGGT  1313

Qy  1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACTTTGACTCCATGCTGGTGAAAATGACC  1260
         |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  1314 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACC  1373

Qy  1261 TGCCGTGGTTCCGACTTTGAAACTGCTTGTCGTGCACGGCGCCGTTGGCTGAGTTC  1320
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  1374 TGCCGTGGTTCCGACTTTGAAACTGCTTGTCGTGCACAGCGCCGTTGGCTGAGTTC  1433

Qy  1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGCATCACCCACCTCCTTCAGCGCT  1380
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1434 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGCATCACCCACCTCCTTCAGCGCT  1493

Qy  1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGCATCACCCGGATTCATTGCCCAGCTTGGCT  1440
         ||||||||||||||||||||||||| |||||||||||||||| |||||||||||||||||
Db  1494 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGCATCACCCGGATTACTTGCCCAGCTTGGCT  1553

Qy  1441 CCACCTGCGGATGATGAGCAGGGACGACCATCCTGGATTACTTGGCAGATGTCACCGTGAAC  1500
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  1554 CCACCTGCTGATGATGAGCAGGGACAGGGAGGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC  1613

Qy  1501 AAGCCTCATGGTGCGTCCAAAGGATGTTGCAGCAATCGATAAAGCTGCCCAACATC  1560
         |||||||||||||||||||||||||||| ||||||||| |||||||||||||||||
Db  1614 AAGCCTCATGGTGCGTCCAAAGGATGTTGCAGCAATCGATAAAGCTGCCTAACATC  1673

Qy  1561 AAGGATCTGCCACTGCCACCCGAGTCCCGTCCACGCGGTTACTGGCCTGAAGCAGCTTGGC  1620
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  1674 AAGGATCTGCCACTGCCACCCGAGTCTTTGTCCGTGGAGGGCTTGGCCCAGCCGCGG  1733

Qy  1621 TTTGCTCGTGATCTCCGTGAGCGACGCACTGGACGTGATCCTACTGACTACTCCGCG  1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1734 TTTGCTCGTGATCTCCGTGAGCGACGCACTGGACGTGATCACTACTACCGCGCG  1793

Qy  1681 GCAACCAGTCTTTTGCTTGCGACCCGAGTCCGCTCATTCGGCACTGAAGCCTGAGACGGCAGAG  1740
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  1794 GCACACCAGTCTTTTGCTTGCGACCCGAGTCCGCTCATTCGGCACTGAAGCCTGCGGCGAGAG  1853

Qy  1741 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCGTGGAGGCCGCGGCGACCTAC  1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1854 GCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCGTGGAGGGCGGCGGCGGCGACCTAC  1913

Qy  1801 GATGTGGCGATGGCGATGGCGTTTCCTCTTTGAGGATCCTGTTGAGGATCCGTGGACAGCGTCGCGCGAG  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1914 GATGTGGCGATGGCGATGGCGTTTCCTCTTTGAGGATCCGTGGACAGCGTCGCGCGAG  1973
```

```
Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1974 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGGATACACC 2033

Qy  1921 CCGTACCCAGACTCCGTCCGCGGCGGCGGTTTGTTGTTAAGGAAGTGCCAGCTCCGGTGGAC 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2034 CCGTACCCAGACTCCGTCCGCGGCGGCGGTTTGTTGTTAAGGAAGTGCCAGCTCCGGTGGAC 2093

Qy  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2094 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA 2153

Qy  2041 GTCCTGGAGACCAACACCGCGCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2154 GTCCTGGAGACCAACACCGCGCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT 2213

Qy  2101 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2214 GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2273

Qy  2161 AAGTCTGGCGCTCACATTCTGGCACTGCGTGTGCTTCGCCCAGCTGCG 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2274 AAGTCTGGCGCTCACATTCTGGCACTGCGTGTGCTTCGCCCAGCTGCG 2333

Qy  2221 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2334 GTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACC 2393

Qy  2281 CACGACACTGCGGGTGCCAGTTGGCTACCTACTTTGCTGCAGCTTCCAGCTTCAAGCTGGTGTGCAGAT 2340
         ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db  2394 CACGACACTGCGGGTGCCAGTTGGCTACCTACTTTGCTGCAGCTTCCAGCTTCAAGCTGGTGTGCAGAT 2453

Qy  2341 GCTGTTGACGGTGCTTCCGCACACTGTCTGCACCACCTCCAGCACCATCCCTGTCTGCC 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2454 GCTGTTGACGGTGCTTCCGCACACTGTCTGCACCACCTCCAGCACCATCCCTGTCTGCC 2513

Qy  2401 ATTGTTGCTGCATTCGCGCACACCCGTCGGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2514 ATTGTTGCTGCATTCGCGCACACCCGTCGGCGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2573

Qy  2461 GACCTCGAGCCGTACTGGGAAGCTGTGGCGCGGACTGCTGTACCTGCCATTGAGTCTCGGAACC 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2574 GACCTCGAGCCGTACTGGGAAGCAGTGGCGCGGACTGCTGTACCTGCCATTGAGTCTCGGAACC 2633

Qy  2521 CCAGGGCCCAACCGGTCGCGTCTACCGCCACGAGAAATCCCAGGCGACAGTTGTCCAACCTG 2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2634 CCAGGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTG 2693

Qy  2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGGATCGCTTCGATCCGATCGTTGCCGAAGACAACTAC 2640
         ||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db  2694 CGTGCACAGGCCACCGCACTGGGCCTTGCTGGATCGCTTCGAACTCATCGAAGACAACTAC 2753

Qy  2641 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2754 GCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2813

Qy  2701 GGGGACCTCGACTCACCTGGTTGCGGGGTGTAGATCCAGCAGACTTTGCTGCCGCAGAC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2814 GGGGACCTCGACTCCACCTGGTTGCGGGGTGTGGATCCAGCAGACTTTGCTGCCGGAT 2873

Qy  2761 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCCGCGGCGGGAGCTTGGTAAC 2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2874 CCACAAAGTACGACATCCCAGACTCTGTCATCGCGTTCCTGCCGCGGCGGGAGCTTGGTAAC 2933

Qy  2821 CCTCCAGGTTGGCTGGCCAGAACCAACCACTGCCGCCACCCCGGCTCGGAAGGCTCCGAAGGC 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2934 CCTCCAGGTTGGCTGGCCAGAGAGCCACTGCCGCCACCCCGGCGCCGCTCCGAAGGC 2993

Qy  2881 AAGGCACCTCTGACGCGAAGTTCCTGAGGAAGACAGAGGCGCACCTCGACGCTGATGATTCC 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2994 AAGGCACCTCTGACGCGAAGTTCCTGAGGAAGACAGAGGCGCACCTCGACGCTGATGATTCC 3053

Qy  2941 AAGGAACGTCGCAACAGCGCTCAACCGCCTGCTGTTCCCGGAAGCCGAAGCCGAAGAGTTCCTC 3000
```

```
Db  3054 AAGGAACGTCGCAATAGCCTCAACCGCCTGCTGTTCCCGGAAGCCAACCGAAGAGTTCCTC 3113

Qy  3001 GAGCACCGTCGCCGCGTTCGGCGAACACACCTCCTGCTGAATGATCGCTGAATTCTTCTACGGA 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3114 GAGCACCGTCGCCGCGTTCGGCGAACACACCTCCTGCTGAATGATCGCTGAATTCTTCTACGGC 3173

Qy  3061 CTGGTCGAGGCGCCGGAGGCCGGAGACTTTGATCCGCTGCCAGATGTGCGCACCCACTGCTTGTT 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3174 CTGGTCGAAGGCGCCGGAGACTTTGATCCGCCAGATGTGCGCACCCCACTGCTTGTT 3233

Qy  3121 CGGCTGGATGCGATCTCTGAGCCAGCAGACGATAAGGGTATGCGCCAACGTGTGGCCAACGTC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3234 CGGCTGGATGCGATCTCTGAGCCAGCAGACGATAAGGGTATGCGCCAACGTGTGGCCAACGTC 3293

Qy  3181 AACGGCCAGATCCGCCCAATGCCGTGTGCGTGACCGCTCCGTTGAGTCTGTGCTTGTCACCGCGACAACC 3240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3294 AACGGCCAGATCCGCCCAATGCCGTGTGCGTGACCGCTCCGTTGAGTCTGTGCTTGTCACCGCGACAACC 3353

Qy  3241 GCAGAAAAGGCAGAATTCCTCCAACAAGGGCCATGTTGTCTGCACCATTCGCTGCGTGTGTGTC 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3354 GCAGAAAAGGCAGAATTCCTCCAACAAGGGCCATGTTGTCTGCACCATTCGCTGCGTGTGTTGTC 3413

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGGCTGGAGATGCAGTGCAGTCGCAATCATCGAG 3360
         ||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3414 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGGCTGGAGATGCAGTGCAGTCGCAATCATCGAG 3473

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCGTTGACGGCGGCGGCAAGATTGAACGCGTTGTG 3420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3474 GCTATGAAGATGGAAGCAACAACAATCACTGCTTCGTTGACGGCGGCGGCAAAATCGATCGGCGTGTG 3533

Qy  3421 GTTCCTGCTGCAACGAACGAAGGTGGCGACTTGATCGTCGTTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3534 GTTCCTGCTGCAACGAACGAAGGTGGCGACTTGATCGTCGTTTCCTAA 3587
```

RESULT 6

```
AAH65730
ID   AAH65730 standard; DNA; 3420 BP.
XX
AC   AAH65730;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 765.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;  99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   P-PSDB; AAG90511.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
```

XX
PS Claim 8; SEQ ID NO: 765; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 3420 BP; 732 A; 1004 C; 939 G; 745 T; 0 other;

Query Match   96.3%; Score 3344.8; DB 22; Length 3420;
Best Local Similarity 98.6%;
Matches 3373; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy   52  GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC  111
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC   60

Qy  112  GGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACCGTAGCT  171
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GGGCGAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACCGTAGCT  120

Qy  172  ATTTACCCCGTGAAGATCGGGGATCATTCCACGCGTCTTTTGCTTCTGAAGCTGTCCGC  231
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  ATTTACCCCGTGAAGATCGGGGATCATTCCACGCGTCTTTTGCTTCTGAAGCTGTCCGC  180

Qy  232  ATTGGTACTGAAGGCTCAAGGCGTACCTGGACATCTGACATCGATGAAATTATCGGTGCA  291
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ATTGGTACCGAAGGCTCAAGGCGTACCTGGACATCTGACATCGATGAAATTATCGGTGCA  240

Qy  292  GCTAAAAAAGTTAAAGCAGATGCTATTTACCCGGGGATATGGCTTCCTGTCTGCTGTCGCC  351
         || |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  241  GCTAAAAAAGTTAAAGCAGATGCTATTTACCCGGGGATACGGCTTCCTGTCTGCTGTCGCC  300

Qy  352  CAGCTTGCCGCCGAGTGCCGCGAAAACGCCATTACTTTTATTGGCCCAACCCCAGAGGTT  411
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  301  CAGCTTGCCGCCGAGTGTCGCGAAAACGCCATTACTTTTATTGGCCCAACCCCAGAGGTT  360

Qy  412  CTTGATCTCACCGGTGATAAGTCTCGTGCGGCAAAACATCGATGACATCGTTAAAGCGCTG  471
         ||||||||||||||||||||||||||| |||||||||||||||| |||||||||||||||
Db  361  CTTGATCTCACCGGTGATAAGTCTCGCGCGGCAAAACATCGATGAGATCGTTAAAGCGCTG  420

Qy  472  GTTTTGGCGGAATCCACCCCATCTTTGTAAAGGCAGTTGCCGGCGGAGCGCGGTATGCGTTT  531
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GTTTTGGCGGAATCCACCCCATCTTTGTAAAGGCAGTTGCCGGCGGAGCGCGGTATGCGTTT  480

Qy  532  CAGACTTACCCCATCTTTGTAAAGGCAGTTGCCGGCGGAGCGCGGTATGCGTTTT  591
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGCGGAGCGCGGTATGCGTTTT  540

Qy  592  GTTTCTTCACCTGACGGTTCGGTATATGTCGAACGTGCTGATTAACCCCCAGCACATTGAA  651
         |||| ||||||||||| |||||||||||||||||||||||||||||||| ||||| |||||
Db  601  GCTTTCGGTTCGGATGGCGCGGTATATGTCGAACGTGCTGATTAACCTCCAGCATATTGAA  660

Qy  652  GCATTCGGGCGACGGTTCGGTATATGTCGAACGTGCTGATTAATGTCGAACGTGCTGATT  711
         ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  601  GCTTTCGGTTCGGATGGCGGTATATGTCGAACGTGCTGATTAATGTCGAACGTGCTGATT  660

Qy  712  GTGTGCAGATCCTTGGCGACTCGCACTGGAAGAGTTGTACACCTTTATGAACGTGACTGCTCA  771
         ||||||||||| ||||||||||||||| ||||||||||||||||||||||| |||||||||||
Db  741  GCGACCTACGATGTGGGCGATGCGGTTCTCTTCTTTGGAGATCGGGGATCGCGGGACGTGACTGCTCA  800

Qy  772  CTGCAGCGGTCGTCACCAAAAAGTTGTCGAACAGCAGCATTGGATCCAGAA  831
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  801  CTGCAGCGGTCGTCACCAAAAAGTTGTCGAACAGCAGCATTGGATCCAGAA  831
```

```
Db  721  CTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAA  780

Qy  832  CTGCGTGATGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCCGCGTCCATTGGTTACCAGGGC  891
         ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  781  CTGCGTGATCGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCCGCCTCCATTGGTTACCAGGGC  840

Qy  892  GCGGGAACCGGTGGAATTCTTGGTCGTCATGAAAAGGGCAACCACGTTTCATCGAAATGAAC  951
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  841  GCGGGAACCCGTGGAATTCTTGGTCGTCATGAAGAAAATGGGCAACCACGTCTTCATCGAAATGAAC  900

Qy  952  CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAGTCGAAGGTGACCGAGGTGACCTTGGTGAAG  1011
         |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  901  CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAGTCGAAGGTGACCGAGGTGACCTTGGTGAAG  960

Qy 1012  GCGCAGATGCGCTTGGCTGCTGCTGGTCGTGGTGCGTCTGTCGTCAGTTGAAGGATTGGTCTGACCCAAGATAAG  1071
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  GCGCAGATGCGCTTGGCTGCTGCTGGTCGTGGTGCGTCTGTCGTCAGTTGAACCTTGAAGGATTGGGTCTGACCCAAGATAAG  1020

Qy 1072  ATCAAGACCACGGTGCAGCACTGCAGTGCCGCATCACCGGTACCGCTCACCAGGCGGAGCTGGCGTTCGT  1131
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  ATCAAGACCACGGTGCAGCACTGCAGTGCCGCATCACCACGAAGATCCAAACAACGGC  1080

Qy 1132  TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT  1191
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 1081  TTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT  1140

Qy 1192  CTTGACGTGCGAGCTCAGCTGCTCCGTGCGGAAATCACCGCACACTTTGACTCCATGCTGGTG  1251
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 1141  CTTGACGTGCGAGCTCAGCTGCTCCGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTG  1200

Qy 1252  AAAATGACCTGCCGTGGTTGCGACTTTGAAACTGCCTGCTGGTTTGCTTGCGTTGCTGCTGCGC  1311
         |||||||||||||||||| ||||||||||| |||||||||||||||| |||||||||||||||
Db 1201  AAAATGACCTGCCGTGGTTGCGACTTTGAAACTGCCTGCTGGTTTGCTTGCGTTGCTGCTGCGC  1260

Qy 1312  GCTGAGTTCACCGGTCTCGTGTTGCAACCAACATTGGTTTCTTGCGTTGCGTTGCTGCGG  1371
         |||||||||||||||||||| |||||||||||||||||| |||||||||||||||||||
Db 1261  GCTGAGTTCACCGGTCTCGTGGTTGCAACCAACATTGGTTTTCTTGCGTTGCGTTGCTGCGG  1320

Qy 1372  GAAGAGACTTCACTTCCAAGCCGCATCGCCACCGGATTTATCGGCGATCACCCACCTC  1431
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1321  GAAGAGACTTCACTTCCAAGCCGCATCGCCACCGCCACCGGATTCATTGGCGATCACCCACCTC  1380

Qy 1432  CTTCAGGCTCCACCTGCCGATGATGAGCGACGGACGACCATCCTGGATTACTTGGCAGATGTC  1491
         ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 1381  CTTCAGGCTCCACCTGCCGATGATGAGCGACGGACGACCATCCTGGATTACTTGGCAGATGTC  1440

Qy 1492  ACCGTGAACAAGCCTGCCACTGCCCGTGTCCCGTGACCGGCCGGTCCGCCGGAAGCAGCTTGGC  1551
         |||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||||
Db 1441  ACCGTGAACAAGCCTAACATCAAGGATCTGCCACTGCCCGTGTCCCGTGACCGGTCCGCCGGAAGCAGCTTGGC  1500

Qy 1552  CCCAACATCAAGGATCTGCCACTGCCCGTGATCTCCGTGAGCGAGACGCACTGGCCAGTTACTGATACCACC  1611
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501  CCTAACATCAAGGATCTGCCACTGCCCGTGATCTCCGTGAGCGAGACGCACTGGCCAGTTACTGATACCACC  1560

Qy 1612  CAGCCGCGTTTGCTCTGGTCGTGCACAGACGCACTGGCCAGTTACTGATACCACC  1671
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561  CCAGCCGCGTTTGCTCTGGTCGTGCACAGACGCACTGGCCAGTTACTGATACCACC  1620

Qy 1672  TTCCGCGATGCACCAGGTCTTGCTTGCGCGACCCGAGTCGCGCTCATTCGCACTGAAGCCT  1731
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1621  TTCCGCGATGCACCAGGTCTTGCTTGCGCGACCCGAGTCCGCGCTCATTCGCACTGAAGCCT  1680

Qy 1732  GCGGCAGAGGCCGTGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCGCTGGGGCGGC  1791
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 1681  GCGGCAGAGGCCGTGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCGCTGGGGCGGC  1740

Qy 1792  GCGACCTACGATGTGGGCGATGCGGTTCTCTTTGGAGGATCGGGGACAGGCTCGACGAG  1851
         |||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db 1741  GCGACCTACGATGTGGGCGATGCGGTTCTCGTTTCCTCTTTGGAGGATCGGGGACAGGCTCGACGAG  1800

Qy 1852  CTGCGCGGAGGCGATGCCGAATGTAAACATTCAGATGCTCGCTTCGCGGCCCGCAAACCGTG  1911
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db 1801  CTGCGCGGAGGCGATGCCGAATGTAAACATTCAGATGCTCGCTTCGCGGCCGCGGCCCGCCAACACCGTG  1860
```

```
Qy   1912 GGATACACCCGTACCCAGACTCCGTCGCCGCGCGGTTTGTTAAGGAAGTTGCCAGCTCC 1971
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GGATACACCCGTACCCAGACTCCGTCGCCGCGCGGTTTGTTAAGGAAGTTGCCAGCTCC 1920

Qy   1972 GGCGTGGACATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 2031
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGCGTGGACATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCA 1980

Qy   2032 ATCGACGCAGTCCTGAGACCAACACCGCGTAGCCGTGGCTATGGCTTATTCTGGT 2091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATCGACGCAGTCCTGAGACCAACACCGCGTAGCCGTGGCTATGGCTTATTCTGGT 2040

Qy   2092 GATCTTCTGATCCAAATGAAAAGCTCTACACCCTTGGATTACTACCTAAAGATGGCAGAG 2151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GATCTTCTGATCCAAATGAAAAGCTCTACACCCTTGGATTACTACCTAAAGATGGCAGAG 2100

Qy   2152 GAGATCGTCAAGTCTCGGCGCTCACATTCTGCCATTAAGGATATGCGTGGTCTGCTTCGC 2211
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   2101 GAGATCGTCAAGTCTCGGCGCTCACATCTTGCCATTAAGGATATGCGTGGTCTGCTTCGC 2160

Qy   2212 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCCGCCGTGAATTCGATCTGCCAGTGCAC 2271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCCGCCGTGAATTCGATCTGCCAGTGCAC 2220

Qy   2272 GTGCACACCCGACACTGCGGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCT 2331
          |||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db   2221 GTGCACACCCGACACTGCGGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCCAAGCT 2280

Qy   2332 GGTGCAGATGCTGTGTGACGGTGCTTCGCACACTGCTGCGCACCACTGCTCCGCCATCC 2391
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGCAGATGCTGTGTGACGGTGCTTCGCACACTGCTGCGCACCACTGCTCCGCCATCC 2340

Qy   2392 CTGTCTGCCATTGTTGCTGCATCCGCACCCGTCGCGCGATACCGGTTTGAGCCTCGAG 2451
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CTGTCTGCCATTGTTGCTGCATCCGCACCCGTCGCGCGATACCGGTTTGAGCCTCGAG 2400

Qy   2452 GCTGTTTCTGACCTCGAGCCGTACTGGAGCGTGTGCGCGGAGCTGTACCTGCCATTTGAG 2511
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GCTGTTTCTGACCTCGAGCCGTACTGGAGCGTGTGCGCGGAGCTGTACCTGCCATTTGAG 2460

Qy   2512 TCTGGAACCCCAGGGCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGGGCAGTTG 2571
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TCTGGAACCCCAGGGCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGGGCAGTTG 2520

Qy   2572 TCCAACCTGCGTGCACAGGCCACCGGACCCGCACTGGGCCTTGCTGTCGCTTCATCGAA 2631
          ||||||||||||||||||||||||||||     ||||||   ||||||||||||||||
Db   2521 TCCAACCTGCGTGCACAGGCCACCGGATCGCGTCGTTCGCGACTCATCGAA 2580

Qy   2632 GACAACTACCAGCCGGTTAATGAGATGCTCGGGACGCCCAACCAAGTCACCCCATCCTCC 2691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GACAACTACCAGCCGGTTAATGAGATGCTCGGGACGCCCAACCAAGTCACCCCATCCTCC 2640

Qy   2692 AAGGTTGTTGCGACCTCGACTCCGACTCCACCTCGTTGGTGCGGGTGTAGATCCAGCAGACTTT 2751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 AAGGTTGTTGCGACCTCGACTCCGACTCCACCTCGTTGGTGCGGGTGTAGATCCAGCAGACTTT 2700

Qy   2752 GCTGCAGACCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGCGTTCCTGCGGCGGAG 2811
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCTGCCAGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGCGCGTTCCTGCGGCGGAG 2760

Qy   2812 CTTGGTAACCTCCAGGTGGCTGGCGTGGCCAGAGAACCACTGCGCACCGGCCGC 2871
          |||||||||||||||||||||||||||||||||||||||||||||||||    |
Db   2761 CTTGGTAACCTCCAGGTGGCTGGCGTGGCCAGAGAACCACTGCGCACCGGAAGGCCGC 2820

Qy   2872 TCCGAGGCAAGGCACCCTTGCTGCGGAAGTTCCTGAGGAAGCAGGCGCACCTCGACGCGT 2931
          ||||| ||||||||||||  |||||||||||||||||||||| ||||||||||||||||
Db   2821 TCCGAAGGCAAGGCACCCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCT 2880

Qy   2932 GATGATTCCAAGGAACGTCGCAACAGCCTCGCAACGCGCACCGA 2991
          |||||||||||||||||||||||| |||||||||||||||||||
Db   2881 GATGATTCCAAGGAACGTCGCAATAGCCTCGCAACGCGCACCGA 2940


Qy   2992 GAGTTCCTGAGCACGTGCGCGCCGCTTCGGCGAACACCTCTGCGCGCTGGATGATCGTGAATTC 3051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GAGTTCCTGAGCACGTGCGCGCCGCTTCGGCGAACACCTCTGCGCGCTGGATGATCGTGAATTC 3000

Qy   3052 TTCTACGGACTGGTCGAGGGCGCCGAGACTTTGATCCGTCCAGATGTGCGCCACCCCA 3111
          ||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||
Db   3001 TTCTACGGCCTGGTCGAAGGCGCCGAGACTTTGATCCGTCCAGATGTGCGCCACCCCA 3060

Qy   3112 CTGCTTGTTCGCTGGATGCGATCTCTGAGCGCAGACGATAAGGGTATGCGCAATGTGTG 3171
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CTGCTTGTTCGCCTGGATGCGATCTCTGAGCGCAGACGATAAGGGTATGCGCAATGTGTG 3120

Qy   3172 GCCAACGTCAACGGCCAGATCCGCCAATGGTGTGTGCAGACCGCTCCGTTGAGTCTGTC 3231
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3121 GCCAACGTCAACGGCCCAGATCCGCCAATGGTGTGTGCAGACCGCTCCGTTGAGTCTGTC 3180

Qy   3232 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCT 3291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 ACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTTGCTGCACCATTCGCT 3240

Qy   3292 GGTGTTGTCACTGTGACTGTGTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3351
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 GGTGTTGTCACCGTGACTGTGTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCA 3300

Qy   3352 ATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGGCAAGATTGAA 3411
          |||||||||||||||||||||  ||||||||||||||||||||||||||||||||    |
Db   3301 ATCATCGAGGCTATGAAGATGAAGCAACAATCACTGCTTCTGTTGACGGGCAAAATCGAT 3360

Qy   3412 CGCGGTTGTGGTTCCTGCTGCAACGAAGGTGGCGGTTGATCGTCGTCGTTTCC 3471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 CGCGGTTGTGGTTCCTGCTGCAACGAAGGTGGCGGTTGATCGTCGTCGTTTCC 3420


RESULT 7
AAF87437
ID   AAF87437 standard; DNA; 4013 BP.
XX
AC   AAF87437;
XX
DT   09-JUL-2001  (first entry)
XX
DE   Corynebacterium thermoaminogenes pc nucleotide sequence.
XX
KW   Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW   thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD;
KW   pdhA; pc; ppc; acn; icd; lpd; odhA; ds.
XX
OS   Corynebacterium thermoaminogenes.
XX
FH   Key             Location/Qualifiers
FT   CDS             319..3738
FT                   /tag= a
FT                   /product= "pc protein"
XX
PN   WO200125447-A1.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-JP06913.
XX
PR   04-OCT-1999;  99JP-0282716.
PR   01-NOV-1999;  99JP-0311147.
PR   21-APR-2000; 2000JP-0120687.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Hirano S,  Nonaka G,  Matsuzaki Y,  Akiyoshi N,  Nakamura K,  Kimura E;
PI   Osumi T,  Matsui K,  Kawahara Y,  Kurahashi O,  Nakamatsu T;
PI   Sugimoto S;
XX
DR   WPI; 2001-300170/31.
DR   P-PSDB; AAB83180.
XX
```

```
PT   Proteins and their DNA useful for microbial production of L-amino acids
PT   -
XX
PS   Claim 34; Page 126-132; 215pp; Japanese.
XX
CC   The present sequence is provided in a specification relating to genes
CC   encoding thermophilic amino acid biosynthesis system enzymes of
CC   the thermotolerant bacterium Corynebacterium thermoaminogenes.
CC   The novel proteins retain at least 10% isocitrate ligase activity
CC   after heating at 50oC for 5 minutes. DNA fragments encoding the
CC   enzymes were isolated from a Corynebacterium thermoaminogenes
CC   chromosomal DNA plasmid library by PCR. The DNA may be used for
CC   developing strains of amino acid producing microorganisms.
XX
SQ   Sequence 4013 BP; 735 A; 1376 C; 1234 G; 668 T; 0 other;

Query Match            64.9%;  Score 2256.2;  DB 22;  Length 4013;
Best Local Similarity  78.7%;  Pred. No. 0;
Matches 2734;  Conservative    0;  Mismatches  733;  Indels    8;  Gaps  3;


Qy     1  GTGACTGTATCACCC--TTGGCGGTCTTCTTGTTGAAAGGAATAATTACTTCCAGTGTCGA    58
          |||||| | ||||||    ||||| |||| | ||| |||||||| |   |||| ||||| ||
Db   269  GTGACTGCCATCACCCAACAGGCTGGTCTCTCGTCGAAAGGAACAAAAACTGTGGTTACAA   328

Qy    59  CTCACACATCTTCAACGCTTCAGCATTCAAAAGATCTTGGTAGCTATTTACC           118
          ||||||| |||  ||||||  ||||||||||| ||||| ||   ||| |||||
Db   329  C---AACACCCTCCACGCTGCCGGCGTTCAAAAAGATCCTGGCCAACCACCGAGGTGAAA   385

Qy   119  TCGGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACC   178
          ||| |||| |||||  |||| || || ||||  ||||| |  ||| ||  |
Db   386  TCGCGGTGCGAGCATTCCCGCGCCTACGACCGGGGCCCAACCGTGGCCATCTACC         445

Qy   179  CCCGTGAAGATCGGGGATCATTCCACCGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTA   238
          ||||  || |||| ||  ||||||||||| |  |||| || ||| | |||| |  |||
Db   446  CCCGGGAGGACCGTGCGTTCCTTCCACCGGTCCTTCGGCCTCCGAGGCGGTGAGGATCGGAA   505

Qy   239  CTGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAA   298
          |  |||||| | || ||||||||  |||| |||| ||||  ||| || |||| | |||
Db   506  CCGAGGCGGTCTCCACCCGTCAAGGCGTCACCCGTCAAGGCGTAGCCTCAAGCGCCGCCAAGA   565

Qy   299  AAGTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTG    358
          |||| |||||| ||||| || |||||| || |||  ||| | |||||| |||| |||||
Db   566  AAGGTGAAAGCCGGACCGCGCGGTCTACCGCGGGGTATGGTTTCCTTTCGGAAAAATGCCCAGCTCG   625

Qy   359  CCCGGCAGTGCGCGGAAAACGGCCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATC   418
          ||||| |||| ||| || ||||| ||   ||| ||   ||||| |||||||| ||| |
Db   626  CGGCGTGAATGCGCGGAGACGCGCGGTTACCTTCCTCATCGGTCCCCACCCCGGAGGTGCTCGACC   685

Qy   419  TCACCGGTGATAAGTCTCGTGCGTAACCGGGATATGCCTCCGCCGGAGAAGGCTGGTCTG   478
          ||||||| || ||||  ||| ||  |  ||| || |||| ||||||||||| ||| || |
Db   686  TCACGGGCGACAAGTCCCAAGCGTCGTGTCCGCCGCAGAGAAGGCCGTGCTCGGCGGTGG    745

Qy   479  CGGAATCCACCCCGAGCAAAAACATCGATGAACATCGTTAAAAAGCGTCGAAGGCCAGACTT   538
          | |||||| ||| ||||    ||||| ||| |||| || | | |||| | | ||| || ||
Db   746  CGGAATCCACCCCCCGAGCACACTGAGTCGATGAGATCGTCAAGAGTGCCCGAGGGCAGAACCT   805

Qy   539  ACCCCATCTTTGTAAAGGCAGTTGCCGCGGTGGTGGCGGACGCGGTATGCGCTTTGTTTCTT   598
          | ||| | ||| |||||||||  ||| ||| | | ||| |||   | |||  || ||| ||
Db   806  ACCCGATCTTCGTCAAGGCCGTCGCAGGCCGTGGTGGGCGGTGGTGTGGCGGTTCGTCGGTCGAGA   865

Qy   599  CACTTGATGAGCTCCCCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCG   658
          ||  |||| | ||||| ||||| ||| | | || | || |||||| || || ||| | ||
Db   866  AGCCCGAGGACCTGCGTGAGCTGCGCAGGAGGAGGCCTCCCGCCGAGGCGGAGGCGGCGGC   925

Qy   659  GCGACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTTGCAGA   718
          ||| | |||  | |  ||  ||| || | |||   || ||| |||| || || | || || 
Db   926  GCGCCGGTCGTCTCACGTCGACGGGCCGTGATCAAACCCCAGCACATCGAGGTGCGCGCAGA   985

Qy   719  TCCTTGGCGATCGCACTGAGAAGTTGTACCTTTATGAACGTGCTCACTGCAGC         778
          ||| |||||| ||  |  || || | | ||     ||  ||  | || ||| ||
Db   986  TCCTCGGTGATCAACCGGACGACGCTTCCTTCCTCGAGGATCGTGCGTTCCCTGCAGC      1045

Qy   779  GTCGTCACCAAAAAGTTGTCGAAATTGCGCACGACGCATTGGATCCAGAACTGCGTG   838
          ||||| | |||||| || | ||||  | | |||| | |||  |  ||||| ||||| 
Db  1046  GCCGCCACCAGAAGTTCGTGGATCGGAGTGGGATGGTGGGATCAGTAAAGCTGGGCG   1105


Qy   839  ATCGCCATTGTGCGGATGCAGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGGCGGGGAA   898
          ||||||||||||| |||||||||| |||  ||||| |||| |||||   ||| || ||| |||
Db  1106  ACCGCATCGTGTCCGATGCCCGTGAGTTCTGCAAATCATCGGATACCAGGGGCCGGCGGCGCA   1165

Qy   899  CCGTGAAATTCTTGTCGATGAAAAGGGCAACCACGTTTCATCGAATGAACCCACGTA   958
          |||| || |||  | | || || ||| || | ||  || || || ||||| ||| |
Db  1166  CCGTGGAGTTCCTCGTCGAGGTCTTCATTGGCCAACCAGTCTTCATTGAGATGAACCCCGCA   1225

Qy   959  TCCAGGTTGAGCACAGCCGGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGA   1018
          |||||| ||| |||  |||||||| | ||||| |  |||  |||| |||  ||||| || |||
Db  1226  TCCAGGTGGAACACACCGGTGACCTGACCTCGCGAGGAGGAGTCACCTCCGACCTGTCAAGGCGCAGA   1285

Qy  1019  TGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGAGGTCTGACCCAAGATAAGATCAAGA   1078
          || | | ||| || |||| |||| |||||||| | || |||| |||  |||| |||| |
Db  1286  TGCACCTGGCCGCCGGGTGCCACCCTGAAGGAACTGGGCCTGACCCCAGGACAAGATCACCA   1345

Qy  1079  CCCACGGTGCAGCACTGCAGTGCCGCCATCACCACGGAAGATCACACCGAGGCTTCGCC   1138
          ||||| |||||| | |  |||| | ||| |||||| |||| | ||  |||| || ||| 
Db  1346  CCCACGGTGTCCCGATCACCACGGCCGGAGACCGTCCACAACTCCAACAACGGCTTCCGGC   1405

Qy  1139  CAGATACCGGAACTATCACCGCTCTACCGGTCTCACCGGCGCTGGCGTTCGTCTTGACG   1198
          | |  | || | || | ||| | |  ||  ||| | |||||| || | |  |  ||| |
Db  1406  CCGACACCGGTGATCATCCGCCGTTCAACCACCGGCTGGCGGGTGTGGCGGTCTCGACG   1465

Qy  1199  GTGCAGCTCAGCTCGGTGGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGA   1258
          ||| || |||| | ||| | || |   || | | || |||||| || | | ||| ||| |
Db  1466  GCGCAGCCCCAGCTCGGCGGCGGGGTCGATCCGCGATCACCGCACATTTCGATTCGGTGTCAAGATGA   1525

Qy  1259  CCTGCGCGGTTCCGACTTTGAAACTGCTGTGTCGTCGTGCACAGCGCGCGCCTTGGCTGAGT   1318
          |||| || || ||| |||  ||| || | | |  |  |||| |  |||| || ||||  |||
Db  1526  CCTGCGCGGTTCCGATTTCGAGACGCGCGTGTCTCCGAGCCCCAGCGGCGCCCTGCTGGTGAGT   1585

Qy  1319  TCACCGTGTCTGGTGTTGCAACCACCATTGGTTTCTTCTGCGGGAAGAGG   1378
          || |||  | | || | || |||| | ||  | |  ||||| |  ||||
Db  1586  TCAACGTCTCCGGCGTGGCGGTGCTGGCGGCTTCCTGCGGCTGCGGGAAGGAAG   1645

Qy  1379  ACTTCACTTCCAAGCGCATCGCCACCGATTTATCGGCGATCACCCACCTCCTTCAGG   1438
          |||||| || |||| | || | |||| | | | ||||  |||| ||| | |  |||
Db  1646  ACTTCACCAAGAGGCCGCATCGACACCGGCTCCATCGGCTCCTCCACCACCTGCTCCAGG   1705

Qy  1439  CTCCACCTGCGGCGATGATGAGCAGGAGGACGCATCGTCTGGATTACTTGCACCGTGA   1498
          || |  | | | ||||||||||||| ||||  ||||||| ||| | |||| |||||||
Db  1706  CCCACCGCCCGACGATGAGCACGGCGGATCCGGGCCGATCGTCGAATACCTGGCGATGTCACCGTGA   1765

Qy  1499  ACAAGCTCATGGTGTCGTCCAAAGGATGTGTCGCACCAATCGATAAGCTGCCCAACA   1558
          ||||   | || | | || ||  ||   ||| ||| | ||||  |||||| ||||||
Db  1766  ACAAACCCCACCGTGAACGCC---CGAGACAGCCCGTCGATAGAGAAGCTGCCCGAGG   1822

Qy  1559  TCAAGGATCTGCCACTGCCAAGCGCCTGGTTCCCGACGCGTTGGGCCTGAAGCGCTG   1618
          ||| | | | | | ||||| ||||| |||| ||||| || |  ||||| || ||||
Db  1823  TGGAGAACATCCGCCTGCCGCCAGCCCCTGGCGTCCGGACCGGCTCGGAAGCAGCTG   1882

Qy  1619  CGTTTGCTGCGTGATCTCCGTGAGAGCCTTTTGTTTCTT         1678
          | ||| | | | | || ||| |||| | || | |   |  ||||  ||| |||| |||
Db  1883  GTTTCGCCCGCCGATCTGCGCGAACAGGATGCCCTGCCGCCGTCACCACCTTCCGGCG   1942

Qy  1679  ATGCACACCAGTCTTTGCTTGCCGCTCGCGAAGTCGTGAAGCTGAAGCGGCCATTCG   1738
          ||||| ||| | ||| || ||| ||  || ||| | || |  | || |||| | |||
Db  1943  ATGCCCACCAGTCCCTGCTGCGCGGCGTGGTGCCACCCCGCGTCGGCGACCCCGGGCAG   2002

Qy  1739  AGGCCGTCGCAAAGCTGACTCCTGACGCTCCTGAGCTTGCAGA          1798
          ||||||| ||||||||| | | | | |  | || || | ||||
Db  2003  GCGCCGTCGCGCGAAAGCTCACCCCCCAGCCAGCCCCAGCACATCGAGGTGTGCCACCT   2062

Qy  1799  ACAATGTGGCCGATGCCTTCCTTCCTTTGAGGATCGTGACGAGTGCGCG   1858
          | | |||||||| | || | |  || |  |||||| | || ||  | ||
Db  2063  ACGACGTGGCCATGGCCATGTCGGACATCCGGAGGATCCGTCTGGGCACGCCGGGCACGTGCTGCTG   2122

Qy  1859  AGGCCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCCACGACCATTGGATCCAGAACTGCGTG   1918
          |||||||| |||||| |  | ||||| | |||||    | | | ||  |||||   ||
Db  1859  GTCGTCACCAAAAAGTTGTCGAAATTGCGCACGACGCATTGGATCCAGAACTGCGTG   838
```

```
Db  2123  AGGCGGATGCCGAATGTGAACATCCAGATGCTGCTGCGTGGTCGCAACACCGTCGGGTACA  2182

Qy  1919  CCCCGTACCCAGACTCCGCTGCCGCGTTGTTAAGGAGAGCTGCCAGCTCCGGCGTGG  1978
          ||||||  ||| || |  || |  |||||||||||
Db  2183  CCCCGTACCCGATTCGGTGTGCGGCCGCGTTTGTGCGGAGGCCGCCAAGTCCGGTGTGG  2242

Qy  1979  ACATCTTCCGATCTTCGACGCGCTTAACGACGTCTCCCAGCAATGCTTCCCAGATGCGTCGACG  2038
          ||||||||||||| || ||||||||||
Db  2243  ACATCTTCCGATCTTCGACGCGCTCAACGACATCTCCCAGCATCGCGCCGGCCATCGACG  2302

Qy  2039  CAGTCTGGAGACCAACACCCGCGGTAGCCGGTGGCTATGGCTTATTCGGTGATCTCT  2098
          |||||||||||||||||||
Db  2303  CCGTCCTGGAGACCGGCACCAGTGTTGCCGAGGTCGCCATGGCGTACTCCGGTGACCTGT  2362

Qy  2099  CTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCG  2158
          ||||| ||||||||||||||||
Db  2363  CCAATCCGGGGGAGAGCTCTACACCCTGACTACTACCTGGCCCGAGCCAGCAGATCG  2422

Qy  2159  TCAAGTCTGCGCTCACATTCTGGCCATTAAGGATATGGCTGTTCGTCGTGTTGCTTCGGCCCAGCTG  2218
          |||||||||||||||||||
Db  2423  TCGACTCCGGTGCACACATCCTGGCCATCAAGGACATGGGCCGCGCCGCGCTCGCCCCG  2482

Qy  2219  CGGTAACCAAGTCGGTGGTGGCCAGTTGGCTACCTACTTTGCTGCACGTGGTGCAG  2278
          ||||||||||||||||
Db  2483  CGGCGCCGCCAAACTGTCGCCGCCCTGCGCGTGCTGCAGCTGCGTGCATGTCCACA  2542

Qy  2279  CCCACGACACTCGCGGTGGCGCCAGTGGCCTGCCACCTTCGGCCGGCCGGGGGCCG  2338
          |||||||||||||||||
Db  2543  CCCACGACACCGCCGGCGGTCAGCTGGCCACCTACCTGGCCGCCCAACCCGGGGCCG  2602

Qy  2339  ATGCTGTTGACGGTGCTTCCGGCACCACTGTTCTCGGGCGTCGCCATCCCTGTGTCTG  2398
          ||||||||||||||||
Db  2603  ATGCCGTCGACGCGCCGGCCTCCGCCACCTGTCCGGTTACCACCTCCCAGCCGTCGATGTCCG  2662

Qy  2399  CCATTGTTGCTGCATTCGCATTCGCATCGCCGATACCGGTTGAGCCTCGAGGCTGTT  2458

Qy  2459  CTGACCTCGAGCCGTACTGGGAAGCGTGTGCGCGGGACGTGTACCGCCATTTGAGTCTGGAA  2518
          |||||||||||||||||
Db  2663  CTCTGGTTGCCGCGTTTGCGCACACCGCACACCGGCGCACCGGCGCCTCAACCTGCAGGCCGTCT  2722

Qy  2519  CCCCAGGGCCCAACCGTCGCGCTCGCCGCTCACGCGCCACGAAATCCCAGGCCGGACAGTTGTCCAACC  2578
          ||||||||||||||||
Db  2723  CCGACCTGGAACCGTACTGGGAGGCGTCCCGCGGGACGTGTACCGCCGTTTGAATCCGGCA  2782

Qy  2579  TGGCTGCACAGGCCACCGCACTGGGCCTTCGATCGCTCGATCGAAGACAACT  2638
          |||||||||||||||||
Db  2783  CCCCGGGCCCGACCACCGGCGCGCGTTTACCGCCCACGAGATCCCCGGCCGTGTCCAACC  2842

Qy  2639  ACGCAGCCGGTTAATGAGATGCTGGACGCGCCCAACCAAGGTCACCCCATCCTCCAAGGTTG  2698
          |||||||||||||||||
Db  2843  TGGCGTGCCCAGGCGCCGTTGCACTGGGTCTGGCCGCACCGGTTCTGAGCTTCATCGAGGACTACT  2902

Qy  2699  TTGGCGACCTCGCACTCCGCACTCCACCTGGTGGTGCGGGTGTAGATCCAGACTTTGCTGCAG  2758
          ||||||||||||||||
Db  2903  ACGCGGCCGGTCAACGAGATGCTGGGTCGTCCGACCAAGGTCACCCCCGTCCTTCCCAAGGTTG  2962

Qy  2759  ACCCACCAAAAGTACGACATCCGACTCTGCGTTCGTCATCGCGGCGGAGCTTGGTA  2818
          |||||||||||||||||
Db  2963  TCGGTGACCTCGCACTCGCACTCCGCTGTGGTGCCGGTGTGCGCCGCCGCCG  3022

Qy  2819  ACCCTCCAGTGGCTGGCCCAGAACCACCTGCGCGACTCCGAAGGCGCTCCGAAG  2878
          |||||||||||||||||
Db  3023  ATCCGGCAGAAGTACGACATCCCCGATTCGGTCATCGCGCACCCCGCCGACTGGGGTA  3082

Qy  2879  GCAAGGCACCTCGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTGCACGCTGATGATT  2938
          |||||||||||||||||
Db  3083  CCCCTTCCCGGTGGCGTGGCTGGCCCAGAACCACCTGCGCGCACCGCTGCTCCAGG  3142

Qy  2939  CCAAGGAACGTCGCAACAGCCTCAACCGCGTTCTCGTTCCCGAAGCCAACCGAAAGAGTTCC  2998
          |||||||||||||||||
Db  3143  GTAAGGCCCCCGTCGGCGGAGATCCCGCGCGCAGGCCCACCACTGGGATTCCGGATGATT  3202

Qy  2999  CCGCGGAGCGTGCGCGGCCAACACCGCGCTCAACCGCGTCGTTCCGCAAGCCGACCGAGAGTTCC  3058
          |||||||||||||||||
Db  3203  CCGCGGAGCGTGCGCGGCCACACCCTCAACCGCGTCGTTCCGCAAGCCGACCGAGAGTTCC  3262

Qy  2999  TCGAGCACCGTCCGCCTTCGGCGCAACACCCTCTGCGCGTGGATGACGTGAATTCTTCTACG  3058
          |||||||||||||||||
Db  3263  TTGAGCACCGTCCGCCTTCGGCGCAACACCCTCCGGATGACCGGAGTTCTTCTACG  3322

Qy  3059  GACTGGTCGAGGGCCGCGGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTG  3118
          |||||||||||||||||
Db  3323  GCTTGAAGGAGGGACGTGAGGAGCTGACTGATCCGACTGGTGTGTCCACCGATGGTGG  3382

Qy  3119  TTCGGCCTGGATGCGATCTCTGAGCCAGACGATAAGGTATGCGCAATGTTGGCCAACG  3178
          |||||||||||||||||
Db  3383  TCCGCCTGGATGCGGTGTCCGAACCGGATGACAAAGGCATGGCGCAACGTGGGTCAACG  3442

Qy  3179  TCAACGGCCGAGATCCGCCAATGCGTCGTGCGGCTCCGTTGAGTCTGTCACCGCAA  3238
          |||||||||||||||||
Db  3443  TCAACGGCCAGATCCGCCCGATCCAAGGTGCGGCGGTTCCGTGGAGTCCGTCACCGCCA  3502

Qy  3239  CCGCAGAAAGCAGATTCCTCCAACAAGGGCATGTTGTCGTCAGCCATTGTGTGTTG  3298
          |||||||||||||||||
Db  3503  CCGGCGGAGAAGCCGATGCCCATGTCGCCCACCATTCGCCGGTGTGG  3562

Qy  3299  TCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCATCGCAATCATCG  3358
          |||||||||||||||||
Db  3563  TCACCGTGACCGTCGCGAGGGTGATGAGATCAAGGCTGGCGACGCGCCATCATTG  3622

Qy  3359  AGGCTATGAAGATGAAGCAACAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTG  3418
          |||||||||||||||||
Db  3623  AGGCCATGAAGATGAAGCAGCCACCATCACCGCGCGTGTCATCGACCGGCCTCG  3682

Qy  3419  TGGTTCCTGCTGCTCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTGTTTCCTA  3473
          |||||||||||||||||
Db  3683  TGGTGCCGCGCGCCACCAAGGTCGAGGGCGCGCACCTCATCGTGGTGTTGCCTA  3737
```

RESULT 8
AAF71419
ID   AAF71419 standard; DNA; 1719 BP.
XX
AC   AAF71419;
XX
DT   30-APR-2001 (first entry)
XX
DE   Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:119.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX
PR   25-JUN-1999;   99US-0141031.
PR   08-JUL-1999;   99DE-1031412.
PR   08-JUL-1999;   99DE-1031413.
PR   08-JUL-1999;   99DE-1031419.
PR   08-JUL-1999;   99DE-1031420.
PR   08-JUL-1999;   99DE-1031424.
PR   08-JUL-1999;   99DE-1031428.
PR   08-JUL-1999;   99DE-1031431.
PR   08-JUL-1999;   99DE-1031433.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031562.
PR   08-JUL-1999;   99DE-1031634.

```
PR   09-JUL-1999;  99DE-1031180.
PR   09-JUL-1999;  99DE-1032227.
PR   09-JUL-1999;  99DE-1032230.
PR   09-JUL-1999;  99US-0143208.
PR   14-JUL-1999;  99DE-1032924.
PR   14-JUL-1999;  99DE-1032973.
PR   14-JUL-1999;  99DE-1033005.
PR   27-AUG-1999;  99DE-1040765.
PR   31-AUG-1999;  99US-0151572.
PR   03-SEP-1999;  99DE-1042076.
PR   03-SEP-1999;  99DE-1042079.
PR   03-SEP-1999;  99DE-1042086.
PR   03-SEP-1999;  99DE-1042087.
PR   03-SEP-1999;  99DE-1042088.
PR   03-SEP-1999;  99DE-1042095.
PR   03-SEP-1999;  99DE-1042123.
PR   03-SEP-1999;  99DE-1042125.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   P-PSDB; AAB79302.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT   metabolism and oxidative phosphorylation protein for production or
PT   modulation of production of fine chemicals e.g. amino acids,
PT   carbohydrates or enzymes -
XX
PS   Claim 3; Page 316-319; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (I), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 1719 BP; 405 A; 457 C; 468 G; 389 T; 0 other;

Query Match           45.0%;  Score 1564.8;  DB 22;  Length 1719;
Best Local Similarity 98.3%;  Pred. No. 0;
Matches  1592;  Conservative  0;  Mismatches  27;  Indels  1;  Gaps  1;
```

```
Qy    241  GAAGGCTCACCAGTCAAGGCTACCTGGACATCGATGAAATTATCGGTCGAGCTAAAAAA  300
                ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    341  GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTCGAGCTAAAAAA  400

Qy    301  GTTAAAGCAGATGCTATTTACCCGGGATATGCTTCCTGTCTGAAAATGCCCAGCTTGCC  360
                ||||||||||||| ||||||||||||||||  |||||||||||| |||||||||||||
Db    401  GTTAAAGCAGATGCCATTTACCCGGGATAACGCTTCCTGTCTGAAAATGCCCAGCTTGCC  460

Qy    361  CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTC  420
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461  CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTC  520

Qy    421  ACCGGTGATAAGTCTCGTGCGGTAACCCGCGAAGAGGCTGGTCTGCCAGTTTTGGCG  480
                ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    521  ACCGGTGATAAGTCTCGCGCGGTAACCCGCGAAGAGGCTGGTCTGCCAGTTTTGGCG  580

Qy    481  GAATCCACCCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC  540
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    581  GAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTAC  640

Qy    541  CCCATCTTTTGTAAAGGCAGTTGCCGGTGGTGGCGACGCGGTATGCGGCTTTGTTTCTTCA  600
                ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    641  CCCATCTTTTGTGAAGGCAGTTGCCGGTGGTGGCGACGCGGTATGCGGCTTTGTTGCTTCA  700

Qy    601  CCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGC  660
                ||||||||||||||||||| |||||||||| |||||||||| |||||||||||||||||||
Db    701  CCTGATGAGCTTCGCAAATTAGCAACAGAACGTGCATCTCGCGAAGCTGAAGCGGCTTTCGGC  760

Qy    661  GACGGTTCGGTATATGTCGAACGTGCTGATTAACCCCAGCACATTGAAGTGCAGATC  720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    761  GATGGCGCGGTATATGTCGAACGTGCTGATTAACCCTCAGCATATTGAAGTGCAGATC  820

Qy    721  CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780
                |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db    821  CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  880

Qy    781  CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCAGCATTGGATCCAGAACTGCGTGAT  840
                |||||||||||||| ||||| |||||||||||||||| ||||||||||||||||||||
Db    881  CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCGTGAT  940

Qy    841  CGCATTTGTGCGGATGCAGTAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC  900
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    941  CGCATTTGTGCGGATGCAGTAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC  1000

Qy    901  GTGGAATTCTTGTGATGAAAAGGCAACCACGTTTCATCGAAATGAACCCACGTATC  960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1001 GTGGAATTCTTGTGATGAAAAGGCAACCACGTCTTCATCGAAATGAACCCACGTATC  1060

Qy    961  CAGGTTGAGCACACCGTGACTGAAGAGTCACCGAGGTGACCTGGTGAAGGCGCAGATG  1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1061 CAGGTTGAGCACACCGTGACTGAAGAGTCACCGAGGTGACCTGGTGAAGGCGCAGATG  1120

Qy    1021 CGCTTGGCTGCGTCGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1121 CGCTTGGCTGCGTCGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1180

Qy    1081 CACGGTGCAGCACTGCAGTGCGCCATCACCACCACGGAGAATCCAAACAACGGCTTCCGCCCA  1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1181 CACGGTGCAGCACTGCAGTGCGCCATCACCACCACGGAGATCCAAACAACGGCTTCCGCCCA  1240

Qy    1141 GATACCGGAACTATCACCGCGTACGGCCGAGCTGGCGTTCGTCTTGACGGT  1200
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1241 GATACCGGAACTATCACCGCGTACGGCCGAGCTGGCGTTCGTCTTGACGGT  1300

Qy    1201 GCAGCTCAGCTCGGTGGCGAAATCACGTGGAAATCACCGGTGAGCACTTTGACTCCATGCTGTGGTGAAAATGACC  1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1301 GCAGCTCAGCTCGGTGGCGAAATCACGGGGCGAAATCACCGGTGAGCACTTTGACTCCATGCTGTGGTGAAAATGACC  1360

Qy    1261 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCGTGCTGCACAGCGCCGCGTTGCGCAGTTC  1320
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1361 TGCCGTGGTTCCGACTTTGAAACTGCTGTTGCGTGCTGCACAGCGCCGCGTTGCGCAGTTC  1420

Qy    1321 ACCGTGTCTGGTGTTGCAACCACATTGGTTTCTTGCGTGCGTGCGTTGCGCGGAAGAGGAC  1380
```

```
Qy      1  GTGACTGCTATCACCCTTGGCGGCTCTCTTGTTGTTGAAAAGGAATAATTACTTCTAGTGTCGACT
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101  GTGACTGCTATCACCCTTGGCGGCTCTCTTGTTGTTGAAAAGGAATAATTACTTCTAGTGTCGACT  160

Qy     61  CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGCGGCAAATC  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161  CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAAACCGCGCGGCAAATC  220

Qy    121  GCGGTCCGTGCTTCCGTGCGCAGCACTGCAGAACCGGTGCGAGCCACGGTAGCTATTTACCCC  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221  GCGGTCCGTGCTTCCGTGCGCAGCACTGCAGAACCGGTGCGAGCCACGGTAGCTATTTACCCC  280

Qy    181  CGTGAAGATCGGGATCATTCCACCCG...  240
                |||||||||||||||||||||||||||
Db    281  CGTGAAGATCGGGATCATTCCACCCG...  340
```

```
Db  1421 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGCTGCGCGGAAGAGGAC 1480

Qy  1381 TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACCTCCTTCAGGCT 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1481 TTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCT 1540

Qy  1441 CCACCTGCGGATGAGCAGGAGGACCCATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1541 CCACCTGCTGATGAGCAGGGACGACGGGATCCTGGATTACTTGGCAGATGTCACCGTGAAC 1600

Qy  1501 AAGCCTCATGGTGTGCGTCCAAAGATGTTGCAGCACCAATCGATAAGCTGCCAACATC 1560
         ||||||||||||||||||||||||||| |||||||| |||||||||||||| |||||
Db  1601 AAGCCTCATGGTGTGCGTCCAAAGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1660

Qy  1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGACCCGCTGAAGCAGCTTGGCCCAGCCGCG 1620
         ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db  1661 AAGGATCTGCCACTGCCACGCGGGTTCCCGTGACCCG-CTGAAGCAGCTTGGCCCAGCCGCG 1719
```

RESULT 9
AAF71420
ID   AAF71420 standard; DNA; 1406 BP.
XX
AC   AAF71420;
XX
DT   30-APR-2001 (first entry)
XX
DE   Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:121.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX

PR   25-JUN-1999;   99US-0141031.
PR   08-JUL-1999;   99DE-1031412.
PR   08-JUL-1999;   99DE-1031413.
PR   08-JUL-1999;   99DE-1031419.
PR   08-JUL-1999;   99DE-1031420.
PR   08-JUL-1999;   99DE-1031424.
PR   08-JUL-1999;   99DE-1031428.
PR   08-JUL-1999;   99DE-1031431.
PR   08-JUL-1999;   99DE-1031433.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031562.
PR   08-JUL-1999;   99DE-1031634.
PR   09-JUL-1999;   99DE-1032180.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032230.
PR   09-JUL-1999;   99US-0143208.
PR   14-JUL-1999;   99DE-1032924.
PR   14-JUL-1999;   99DE-1032973.
PR   14-JUL-1999;   99DE-1033005.
PR   27-AUG-1999;   99DE-1040765.
PR   31-AUG-1999;   99US-0151572.
PR   03-SEP-1999;   99DE-1042076.
PR   03-SEP-1999;   99DE-1042079.
PR   03-SEP-1999;   99DE-1042086.
PR   03-SEP-1999;   99DE-1042087.
PR   03-SEP-1999;   99DE-1042088.
PR   03-SEP-1999;   99DE-1042095.
PR   03-SEP-1999;   99DE-1042123.
PR   03-SEP-1999;   99DE-1042125.
XX
PA   (BADI) BASF AG.
XX
PI   Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   P-PSDB; AAB79303.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT   metabolism and oxidative phosphorylation protein for production or
PT   modulation of production of fine chemicals e.g. amino acids,
PT   carbohydrates or enzymes -
XX
PS   Claim 3; Page 321-323; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (II), (III), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 1406 BP; 325 A; 381 C; 385 G; 315 T; 0 other;

Query Match   38.9%; Score 1351.8; DB 22; Length 1406;
Best Local Similarity 98.0%; Pred. NO. 0;
Matches 1379; Conservative 0; Mismatches 27; Indels 1; Gaps   1;

```
Qy  214 GCTTCTGAAGCTGTCCCATTGGTACTGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 273
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCTTCTGAAGCTGTCCCATTGGTACTGAAGGCTCACCAGTCAAGGCGTACCTGGACATC 60

Qy  274 GATGAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCCGGGGATATGGC 333
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GATGAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGGATACGGC 120

Qy  334 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGCGGCGGAAAACCGGCATTACTTTATT 393
        ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  121 TTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGCGGAAAACCGGCATTTACTTTATT 180

Qy  394 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTAACCGCCGCG 453
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCG 240

Qy  454 AAGAAGGCTGGTCTGCCAGTTTTGGGGGAATCCACCCCCGAGCAAAAACATCGATGACATC 513
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAGAAGGCTGGTCTGCCAGTTTTGGGGGAATCCACCCCCGAGCAAAAACATCGATGAGATC 300

Qy  514 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 573
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCGGTGGTGGC 360

Qy  574 GGACGCGGTATGCGCTTGTTTTCTTCACCTGATGAGCTCCGCAAATTGGCAACAGAAGCA 633
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GGACGCGGTATGCGCTTGTTTTGTTCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCA 420

Qy  634 TCTCGTGAAGCTGAAGCGGCATTCGGCGACGTTCGGTATATGTGCTGTGATT 693
        ||| |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    421  TCTCGTGAAGCTGAAGCGGCTTTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATT  480
                ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    694  AACCCCCAGCACATTGAAGTGCAGATCCTTGGGCGATCGCCACTGGAGAGTTGTACACCTT  753

Db    481  AACCCTCAGCATATGAAGTGCAGATCCTTGGCGATCACACTGGAGAGTTGTACACCTT     540
                ||||||||||||||||||||||||||||||||||||||||||||||
Qy    754  TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGCGAAATTGCGCCAGCA  813

Db    541  TATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGCGAAATTGCGCCAGCA  600
                ||||||||||||||||||||||||||||||||||||||||||||||
Qy    814  CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCCGGATGCAGTAAAGTTCTGCCGC  873

Db    601  CAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCCGGATGCAGTAAAGTTCTGCCGC  660
                ||||||||||||||||||||||||||||||||||||||||||||||
Qy    874  TCCATTGGTTACCAGGGCGCGGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC  933

Db    661  TCCATTGGTTACCAGGGCGCGGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCAC  720
                ||||||||||||||||||||||||||||||||||||||||||||||
Qy    934  GTTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACC  993

Db    721  GTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGGTGCAACCTTGAAGGAATTG  780
                |||||||||||||||||||||||||||||||||||||
Qy    994  GAGGTGGACCTGGTTGAAGGCGCAGATGCCGTTGGCTTGGTGCTGCTTGGTGCAACCTTGAAGGAATTG  1053

Db    781  GAGGTGGACCTGGTTGAAGGCGCAGATGCGCTTGGCTGTGCTGCTTGGTGCAACCTTGAAGGAATTG  840
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1054 GGTCTGACCCAAGATAAGATCAAGACCCCACGGTGCAGCACTGCGAGCTGCCGCATCACCACG  1113

Db    841  GGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCGAGCTGCCGCATCACCACG  900
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1114 GAAGATCCAAACAACACGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA  1173

Db    901  GAAGATCCAAACAACACGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCA  960
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1174 GGCGGAGCTCGGCGTTCGTCTTGTCTTGAAAAATGACCTGCCGTTCCGGTTGTTCCGTTGCT  1233

Db    961  GGCGGAGCTCGGCGTTCGTCTTGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACAC  1020
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1234 TTTGACTCCATGCTGGTGAAAAATGACCTGCCGTTCAGCTTGAAACTGCTGTTGCT  1293

Db    1021 TTTGACTCCATGCTGGTGAAAAATGACCTGCCGTTCAGCTTGAAACTGCTGTTGCT  1080
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1294 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTC  1353

Db    1081 CGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCGTGTTGCAACCAACATTGGTTTC  1140
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1354 TTGGCTGCGTTGCTGCGCGGAAGAGACTTCACTTCCAAGCGCATCGCCACCGGATTTATC  1413

Db    1141 TTGGCTGCGTTGCTGCGCGGAAGAGACTTCACTTCCAAGCGCATCGCCACCGGATTCATT  1200
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1414 GGGGATCACCCACCACCTCCTTCAGGCTCCACCTGCCGGATGATGAGCAGCGGACGCATCCTG  1473

Db    1201 GCCGATCACCCGCCACCCGCACCTCCTTCAGGCTCCACCTGCCGGATGATGAGCAGCGGACGCATCCTG  1260
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1474 GATTACTTGGCAGATGTCACCGTGAACACAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA  1533

Db    1261 GATTACTTGGCAGATGTCACCGTGAACACAGCCTCATGGTGTGCGTCCAAAGGATGTTGCA  1320
                ||||||||||||||||||||||||||||||||||||||||||||
Qy    1534 GCACCAATGCGATAAGCTGCCCAACATCAAGGATCTGCCACCGCCGGTTCCGTGAC  1593

Db    1321 GCTCCTATCGATAAGCTGCCTAACATCAAGGATCTGCCACTGCCCACGCGTTCCCGTGAC  1380
                ||||||||||||||||||||||||||||||||||||
Qy    1594 CGCCTGAAGCAGCTTGGCCCAGCCGCG  1620

Db    1381 CG-CTGAAGCAGCTTGGCCCAGCCGCG  1406
```

```
RESULT 10
AAF71417
ID   AAF71417 standard; DNA; 1083 BP.
XX
AC   AAF71417;
XX
```

CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 1083 BP; 211 A; 343 C; 300 G; 229 T; 0 other;

Query Match   29.4%;   Score 1020.2; DB 22; Length 1083;
Best Local Similarity 98.4%;   Pred. No. 1.2e-268;
Matches 1062;   Conservative   0; Mismatches 13; Indels 4; Gaps 3;

Qy 1597 CTGAAGCAGCTTGGCCCAGCGCGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCA 1656
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5 CTGAAGCAGCTTGGCCCAGCGCGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCA 64

Qy 1657 GTTACTGATACCACCTTCCCGGATGCACACCAGTCTTTGCTTGCCGACCCGAGTCCGCTCA 1716
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 GTTACTGATACCACCTTCCCGGATGCACACCAGTCTTTGCTTGCCGACCCGAGTCCGCTCA 124

Qy 1717 TTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTG 1776
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125 TTCGCACTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTG 184

Qy 1777 GAGGCCTGGGGCGGCCGCACCTACGATGTGGCGATGCGGTTCCTCTTGAGGATCCGTGG 1836
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 185 GAGGCCTGGGGCGGCCGCACCTACGATGTGGCGATGCGGTTTCCTCTTTGAGGATCCGTGG 244

Qy 1837 GACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGC 1896
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 245 GACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGC 304

Qy 1897 GGCCGCCAAACACCGGTGGAGAGGATGTGTCAAGTCTGCGGCGCTCAAATTCTGCGCCGTTTGTTAAG 1956
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 305 GGCCGCCAAACACCGGTGGGAGAGGATGTGTCAAGTCTGCGGCGCTCAACATTCTGCGCCGTTTGTTAAG 364

Qy 1957 GAAGCTGCCAGCTCCGCGTGACATCTTCCGACGCGCTTAACGACGTCTCC 2016
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 365 GAAGCTGCCAGCTCCGCGTGACATCTTCCGACGCGCTTAACGACGTCTCC 424

Qy 2017 CAGATGCGTCCAGCAATGACGACGCAGTCGCGGACCAACACCGCGGTAGCCGAGGTGGCT 2076
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 425 CAGATGCGTCCAGCAATGACGACGCAGTCGCGGACCAACACCGCGGTAGCCGAGGTGGCT 484

Qy 2077 ATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTTACACCCTGGATTACTAC 2136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485 ATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTTACACCCCTGGATTACTAC 544

Qy 2137 CTAAAGATGGCAGAGAGATCGTCAAGTCTGCGGCTCACATTCTGCCATTAAGGATATG 2196
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 545 CTAAAGATGGCAGAGAGATCGTCAAGTCTGCGGCTCACATCTTGCCATTAAGGATATG 604

Qy 2197 GCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCGCCGTGAATTC 2256
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 605 GCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCGCCGTGAATTC 664

Qy 2257 GATCTGCCAGTCAGTGCAGATGCGTGTGACGGTGCTTCC-GCACCACTGTCTGGCAC 2316
         ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 665 GATCTGCCAGTCAGTGCACGTGCACACCCAGCTGCGGTGCTTCCGGCACCACTGTCTGGCAC 724

Qy 2317 GCTGCAGTCAAGCTGGTGCAGATGCGTGTGACGGTGCTTCC-GCACCACTGTCTGGCAC 2375
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 725 GCTGCAGTCAAGCTGGTGCAGATGCGTGTGACGGTGCTTCCGGCACCACTGTCTGGCAC 784

Qy 2376 CACCTCCCA--GCCATCCCGTGCTGCCATTGTTGCTGCCATTGTTGCTGCGGAT 2433
         |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 785 CACCTCCCAAGCCATCCCGTGCTGCCATTGTTGCTGCCATTCGGCGCACCCGGTCGCGGAT 844

Qy 2434 ACCGGTTTGAGCCTCGAGGCTGTGTTTCTGACCTCGAGCCGTACTGGGAAGCTGTGCGCGGA 2493
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 845 ACCGGTTTGAGCCTCGAGGCTGTGTTTCTGACCTCGAGCCGTACTGGGGAAGCAGTGCGCGGGA 904

Qy 2494 CTGTACCTGCCATTGAGTCTCGGAACCCCAGCGCCCAACCGGTCGCGTCTCACCGGCCACGAA 2553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 905 CTGTACCTGCCATTGAGTCTCGGAACCCCAGCGCCCAACCGGTCGCGTCTCACCGGCCACGAA 964

Qy 2554 ATCCCAGGCGGACAGTTGTCCAACCTGGTGCACAGCGCCACCGCACTGGGCCTTGCTGAT 2613
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 965 ATCCCAGGCGGACAGTTGTCCAACCTGGTGCACAGCGCCACCGCACTGGGCCTTGCCGGAT 1024

Qy 2614 CGCTTCGAGCTCATCGAAGACAACTACGC-AGCCGTTAATGAGATGCTGGGACGCCCCAA 2671
         |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 1025 CGTTTCGAACTCATCGAAGACACACTACGCACGCAGCCGTTAATGAGATGCTGGGGACGCCCCAA 1083

RESULT 11
AAF71418
ID   AAF71418 standard; DNA; 1083 BP.
XX
AC   AAF71418;
XX
DT   30-APR-2001 (first entry)
XX
DE   Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:117.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX
PR   25-JUN-1999; 99US-0141031.
PR   08-JUL-1999; 99DE-1031412.
PR   08-JUL-1999; 99DE-1031413.
PR   08-JUL-1999; 99DE-1031419.
PR   08-JUL-1999; 99DE-1031420.
PR   08-JUL-1999; 99DE-1031424.
PR   08-JUL-1999; 99DE-1031428.
PR   08-JUL-1999; 99DE-1031431.
PR   08-JUL-1999; 99DE-1031433.
PR   08-JUL-1999; 99DE-1031434.
PR   08-JUL-1999; 99DE-1031510.
PR   08-JUL-1999; 99DE-1031562.
PR   08-JUL-1999; 99DE-1031634.
PR   09-JUL-1999; 99DE-1032180.
PR   09-JUL-1999; 99DE-1032227.
PR   09-JUL-1999; 99DE-1032230.
PR   09-JUL-1999; 99US-0143208.
PR   14-JUL-1999; 99DE-1032924.
PR   14-JUL-1999; 99DE-1032973.
PR   14-JUL-1999; 99DE-1033005.
PR   27-AUG-1999; 99DE-1040765.
PR   31-AUG-1999; 99US-0151572.
PR   03-SEP-1999; 99DE-1042076.
PR   03-SEP-1999; 99DE-1042079.
PR   03-SEP-1999; 99DE-1042086.
PR   03-SEP-1999; 99DE-1042087.
PR   03-SEP-1999; 99DE-1042088.
PR   03-SEP-1999; 99DE-1042095.
PR   03-SEP-1999; 99DE-1042123.
PR   03-SEP-1999; 99DE-1042125.

```
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   P-PSDB; AAB79301.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT   metabolism and oxidative phosphorylation protein for production or
PT   modulation of production of fine chemicals e.g. amino acids,
PT   carbohydrates or enzymes -
XX
PS   Claim 3; Page 314-315; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 1083 BP; 211 A; 343 C; 300 G; 229 T; 0 other;

Query Match        29.4%;  Score 1020.2; DB 22; Length 1083;
Best Local Similarity 98.4%;  Pred. No. 1.2e-268;
Matches 1062;  Conservative  0;  Mismatches 13;  Indels  4;  Gaps  3;

Qy   1597 CTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCA 1656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 CTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCA 64

Qy   1657 GTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCCTTGCTTGCCGAGTCCGCTCA 1716
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 GTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCTGCACCCGAGTCCGCTCA 124

Qy   1717 TTCGCCACTGAAGCCTGCGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTG 1776
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    125 TTCGCCACTGAAGCCTGCGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTG 184

Qy   1777 GAGGCCTGGGGCGGCGGCGACCTACGAGTGGCGACGATGCGTTCCTCCTTTGAGGATCCGTGG 1836
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    185 GAGGCCTGGGGCGGCGGCGACCTACGAGTGGCGACGATGCGTTCTCTTTGAGGATCCGTGG 244

Qy   1837 GACAGGCTGCGACGAGCTGCGCGAGGCGATGCCGAGGCGAATGTAAACATTCAGATGCTGCTTCGC 1896
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    245 GACAGGCTGCGACGAGCTGCGCGAGGCGATGCCGAGGCGAATGTAAACATTCAGATGCTGCTTCGC 304

Qy   1897 GGCGCAACACCGTGGGATACACCCGTACCCAGACTCCGTCTGCCGGCGCGGCGTTTGTTAAG 1956
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    305 GGCGCAACACCGTGGGATACACCCGTACCCAGACTCCGTCTGCCGGCGCGGCGTTTGTTAAG 364

Qy   1957 GAAGCTGCCAGCTCCGGCTGGACATCTTCCGACATCTTCGACGCGGCTTAACGACGTCTCC 2016
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    365 GAAGCTGCCAGCTCCGGCTGGACATCTTCGACATCTTCGACGCGGCTTAACGACGTCTCC 424

Qy   2017 CAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCT 2076
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    425 CAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCT 484

Qy   2077 ATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTAC 2136
             |||||||||||||||||||||||||||||||||||||||||
```

```
Db    485 ATGGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTAC 544

Qy   2137 CTAAAGATGGCAGAGGAGATCGTCAAGTCTGCGGCTCACATTCTGGCCATTAAGGATATG 2196
             |||||||||||||||||||||||||||||||||||||||||||||||
Db    545 CTAAAGATGGCAGAGGAGATCGTCAAGTCTGCGGCTCACATTCTGGCCATTAAGGATATG 604

Qy   2197 GCTGGTCTGCTTCGCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCGCCGTGAATTC 2256
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    605 GCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCGCCGTGAATTC 664

Qy   2257 GATCTGCCAGTGCAGTGCACGTGCACACCCAGCACACTGCGGTGGCCAGTTGGCTACCTACTTT 2316
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    665 GATCTGCCAGTGCAGTGCACGTGCACACCCAGCACACTGCGGTGGCCAGTTGGCAACCTACTTT 724

Qy   2317 GCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTGCTTCC-GCACCACTGTCTGGCAC 2375
             ||||||||||||||||||||||||||||||||||||| |||||||
Db    725 GCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTGCTTCCGGACCACTGTCTGGCAC 784

Qy   2376 CACCTCCCA--GCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 2433
             ||||||||  |||||||||||||||||||||||||||||||
Db    785 CACCTCCCAAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGAT 844

Qy   2434 ACCGGTTTGAGCCTCGAGGCGTGTTTCTGACCTCGACCGCGTACTGGGAAGCTGTGCGCGGA 2493
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    845 ACCGGTTTGAGCCTCGAGGCGTGTTTCTGACCTCGACCGCGTACTGGGAAGCAGTGCGCGGA 904

Qy   2494 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCCGGTCGCGTCTACCGCCACGAA 2553
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    905 CTGTACCTGCCATTTGAGTCTGGAACCCCAGGCCCAACCCGGTCGCGTCTACCGCCACGAA 964

Qy   2554 ATCCCAGGCGACAGTTGTCCAACCTGCGTGCACAGGCACCGCACCGCTTGCTGTGAT 2613
             ||||||||||||||||||||||||||||||||||||||||||||||
Db    965 ATCCCAGGCGACAGTTGTCCAACCTGCGTGCACAGGCACCGCACCGCTTGCTGCGGAT 1024

Qy   2614 CGCTTCGAGCTCATCGAAGACAACAACTACGC-AGCCGTTAATGAGATGCTGGACGCCCAA 2671
             ||||||||||||||||||||||||||||||| |||||||||||
Db   1025 CGTTTCGAACTCATCGAAGACAACAACTACGCAAGCCGTTAATGAGATGCTGGGACGCCCAA 1083
```

```
RESULT 12
AAF71415
ID   AAF71415 standard; DNA; 939 BP.
XX
AC   AAF71415;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:111.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX
PR   25-JUN-1999;      99US-0141031.
PR   08-JUL-1999;      99DE-1031412.
PR   08-JUL-1999;      99DE-1031413.
PR   08-JUL-1999;      99DE-1031419.
PR   08-JUL-1999;      99DE-1031420.
PR   08-JUL-1999;      99DE-1031424.
PR   08-JUL-1999;      99DE-1031428.
```

```
PR   08-JUL-1999;   99DE-1031431.
PR   08-JUL-1999;   99DE-1031433.
PR   08-JUL-1999;   99DE-1031434.
PR   08-JUL-1999;   99DE-1031510.
PR   08-JUL-1999;   99DE-1031562.
PR   08-JUL-1999;   99DE-1031634.
PR   09-JUL-1999;   99DE-1032180.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032230.
PR   09-JUL-1999;   99US-0143208.
PR   14-JUL-1999;   99DE-1032924.
PR   14-JUL-1999;   99DE-1032973.
PR   14-JUL-1999;   99DE-1033005.
PR   27-AUG-1999;   99DE-1040765.
PR   31-AUG-1999;   99US-0151572.
PR   03-SEP-1999;   99DE-1042076.
PR   03-SEP-1999;   99DE-1042079.
PR   03-SEP-1999;   99DE-1042085.
PR   03-SEP-1999;   99DE-1042087.
PR   03-SEP-1999;   99DE-1042088.
PR   03-SEP-1999;   99DE-1042095.
PR   03-SEP-1999;   99DE-1042123.
PR   03-SEP-1999;   99DE-1042125.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   P-PSDB; AAB79298.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
     metabolism and oxidative phosphorylation protein for production or
     modulation of production of fine chemicals e.g. amino acids,
     carbohydrates or enzymes -
XX
PS   Claim 3; Page 306-307; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (II), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 939 BP; 202 A; 283 C; 261 G; 193 T; 0 other;

Query Match             25.4%;  Score 883;  DB 22;  Length 939;
Best Local Similarity   98.3%;  Pred. No. 3.9e-231;
  Matches  903;  Conservative    0;  Mismatches   15;  Indels    1;  Gaps    1;

Qy   2557 CCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACCTGGGCCTTGCTGATCGC 2616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      1 CCAGGCGGACAGTTGTCCAACCTGCGCACCGCACCGCACCTGGGCCTTGCTGGATCGT 60
Qy   2617 TTCGAGCTCATCGAAGACAACTACGC-AGCCGTTAATGAGATGCTGGGACGCCCAACCAA 2675
          |||||.|||||||||||||||||||| |||||||||||||||||| |||||||||||||||
Db     61 TTCGAACTCATCGAAGACAACTACGCAAGCCGTTAATGAGATGCTCGGGACGCCCAACCAA 120
Qy   2676 GGTCACCCCATCCTCCAAGGTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGGTGT 2735
          |||||||.|||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GGTCACCCATCCTCCAAGGTGTTGGCGACCTCGCACTCCACCTCGTTGGTGCGGGTGT 180
Qy   2736 AGATCCAGACAGACTTTCGTCAGAGCACCCACAAAAGTACGACATCCCAGACTCTGTCATCGC 2795
          |||||||||||||||.||.|||||||.||.||||||||||||||||||||||||||||||||
Db    181 GGATCCAGACAGACTTTGCTGCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGC 240
Qy   2796 GTTCCTGCGCGGCGAGCTTGGTAACCTCCAGTGGCTGGCCAGAACCACCTCGCGCACCCG 2855
          ||||||||||||||||||||||||||||||||||||||||||||.|||||||||||||||
Db    241 GTTCCTGCGCGGCGAGCTTGGTAACCTCCAGTGGCTGGCCAGAGCCACCTCGCGCACCCG 300
Qy   2856 CGCACTCGAAGGCCGTCCGAAGGCAAGGCACCTCTGACGAAGTTCCTGAGGAAGAGCA 2915
          |||||||.|||||||.|||||||||||||||||||||||.|||||||||||||||||||
Db    301 CGCACTCGGAAGGCCGCCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCA 360
Qy   2916 GGGGCACCTCGACGCTGATGATTCCAAGGAAGTCGCAACAGCCTCAACCGCCTCTGTT 2975
          ||.|||||||||||||||||||||||||||||.||||.|||||||||||||||||||||
Db    361 GGCGCACCTCGACGCTGATGATTCCAAGGAAGCTGCAATAGCCTCAACCGCCTCTGTT 420
Qy   2976 CCCGAAGCCAACCGAAGAGTTCCTGACCACCGCCGTTCGGCAACACCTCTGCCCT 3035
          |||||||.|||||||||||||||||||||||||||||||||||||||||||.||||
Db    421 CCCGAAGCAACCGAAGAGTTCCTGACCACCGCCGTTCGGCAACACCTCTGCCT 480
Qy   3036 GGATGATCGTGAATTCTTCTACGGACTGGTCGAGGGCGCCGGAGACTTTGATCCGCCTGCC 3095
          ||||||||||||||||||||||||.||||||||||.|||||||||||||||||||||.|||
Db    481 GGATGATCGTGAATTCTTCTACGGCCTGGTCGAGAGCCGCGGAGACTTTGATCCGCCCTGCC 540
Qy   3096 AGATGTGCGCACCCCACTGCTTGTTCCTGATGCGATCTCTGAGCGCAGCAGAGATAAGGG 3155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGATGTGCGCACCCCACTGCTTGTTCCTGATGCGATCTCTGAGCGCAGCAGAGATAAGGG 600
Qy   3156 TATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCG 3215
          ||||||||||||||||.|||||||||||||||||||||||||||||||||||||||||||
Db    601 TATGCGCAATGTTGTGCGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCG 660
Qy   3216 CTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGCCATGT 3275
          |||||||||||||||||||||||||||||||||||||||||||||||||||||.|||||
Db    661 CTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGT 720
Qy   3276 TGCTGCACCATTGCGTCGGTGTGTTGTCACTGTGCTGTGTCGAAGGTGATGAGTCAGGC 3335
          |||||||||||||||||||||||||||||||.||.|||||||||||||||||||||||||
Db    721 TGCTGCACCATTGCGTCGGTGTGTTGTCACCGTGACGTGTTGCTGAAGGTGATGAGTCAGGC 780
Qy   3336 TGGAGATGCAGTCGCAATCATCGAGATGTCGAAGATGGAAGCAACAATCACTGCTTCTGT 3395
          ||||||||||||||||||||||||.||.||.||||||||||||||||||||||||||||||
Db    781 TGGAGATGCAGTCGCAATCATCGAGGTCATCGAGGTATGAAGATGGAAGCAACAATCACTGCTTCTGT 840
Qy   3396 TGAGGCCAAGATTGAACGCGTTGGTGTTCCTGCTGCAACGAAGGTGGCGACTT 3455
          |||||||||||||||||||||||||.|||||||||||||||||||||||||||
Db    841 TGACGGCAAAATCGATCGCGTTGGTGTTCCTGCTGCAACGAAGGTGGCGACTT 900
Qy   3456 GATCGTCGTTCGTTCCTAA 3474
          |||||||||||||||||||
Db    901 GATCGTCGTTCGTTCCTAA 919
```

RESULT 13

```
AAF71416
ID   AAF71416 standard; DNA; 939 BP.
XX
AC   AAF71416;
XX
DT   30-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:113.
XX
KW   Corynebacterium glutamicum; carbon metabolism and energy production;
KW   SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW   fine chemical production; organic acid; proteinogenic amino acid;
KW   nonproteinogenic amino acid; organic acid; purine base; pyrimidine base; nucleoside;
KW   nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW   carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KW   diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
```

```
OS   Corynebacterium glutamicum.
XX
PN   WO200100844-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00943.
XX
PR   25-JUN-1999; 99US-0141031.
PR   08-JUL-1999; 99DE-1031412.
PR   08-JUL-1999; 99DE-1031413.
PR   08-JUL-1999; 99DE-1031419.
PR   08-JUL-1999; 99DE-1031420.
PR   08-JUL-1999; 99DE-1031424.
PR   08-JUL-1999; 99DE-1031428.
PR   08-JUL-1999; 99DE-1031431.
PR   08-JUL-1999; 99DE-1031433.
PR   08-JUL-1999; 99DE-1031434.
PR   08-JUL-1999; 99DE-1031510.
PR   08-JUL-1999; 99DE-1031562.
PR   08-JUL-1999; 99DE-1031634.
PR   09-JUL-1999; 99DE-1032180.
PR   09-JUL-1999; 99DE-1032227.
PR   09-JUL-1999; 99DE-1032230.
PR   09-JUL-1999; 99US-0143208.
PR   14-JUL-1999; 99DE-1032924.
PR   14-JUL-1999; 99DE-1032973.
PR   14-JUL-1999; 99DE-1033005.
PR   27-AUG-1999; 99DE-1040765.
PR   31-AUG-1999; 99US-0151572.
PR   03-SEP-1999; 99DE-1042076.
PR   03-SEP-1999; 99DE-1042079.
PR   03-SEP-1999; 99DE-1042086.
PR   03-SEP-1999; 99DE-1042087.
PR   03-SEP-1999; 99DE-1042088.
PR   03-SEP-1999; 99DE-1042095.
PR   03-SEP-1999; 99DE-1042123.
PR   03-SEP-1999; 99DE-1042125.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR   WPI; 2001-061975/07.
DR   P-PSDB; AAB79299.
XX
PT   New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT   metabolism and oxidative phosphorylation protein for production or
PT   modulation of production of fine chemicals e.g. amino acids,
PT   carbohydrates or enzymes -
XX
PS   Claim 3; Page 308-310; 1246pp; English.
XX
CC   AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC   metabolism and oxidative phosphorylation (SMP) proteins given in
CC   AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC   energy production. The C. glutamicum SMP gene can be used in vectors
CC   (II) for expression in host cells and production or modulation of
CC   production of fine chemicals, such as, an organic acid, a proteinogenic
CC   or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC   a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC   acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC   cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC   (III) encoded by them are used for diagnosing the presence or activity of
CC   Corynebacterium diphtheriae in a subject. (I), (III) or host cells
CC   containing them are used to map genomes of organisms related to
CC   C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC   in evolutionary studies, in determining SMP protein regions required
CC   for function, in modulating SMP protein activity, in modulating the
CC   metabolism of sugars, and in modulating high-energy molecule production
CC   in a cell (i.e. ATP, NADPH).
XX
SQ   Sequence 939 BP; 202 A; 283 C; 261 G; 193 T; 0 other;
```

```
Query Match             25.4%;  Score 883;  DB 22;  Length 939;
Best Local Similarity   98.3%;  Pred. No. 3.9e-231;
Matches  903;  Conservative  0;  Mismatches  15;  Indels  1;  Gaps  1;

Qy   2557 CCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGC 2616
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      1 CCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATCGT 60

Qy   2617 TTCGAGCTCATCGAAGACAACTACGC-AGCCGTTAATGAGATGCTGGACGCCCAACCAA 2675
          |||||| |||||||||||||||| ||| |||||||||||||||||||||||||||||||
Db     61 TTCGAACTCATCGAAGACAACTACGCCAAGGCCGTTAATGAGATGCTGGACGCCCAACCAA 120

Qy   2676 GGTCACCCCATCCTCCAAGGTGTTGGCGACCTCGCACTCCACCTGGTTGGTCGCGGTGT 2735
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    121 GGTCACCCCATCCTCCAAGGTGTTGGCGACCTCGCACTCCACCTCGTTGGTCGCGGTGT 180

Qy   2736 AGATCCAGCAGACTTTGCTGCGAGACCCACAAAAGTACGACATCCCAGACTCTGTCATCGC 2795
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGATCCAGCAGACTTTGCTGCGATCCACAAAAGTACGACATCCCAGACTCTGTCATCGC 240

Qy   2796 GTTCCTGCGGCGGGAGCTTGGTAACCCTCCAGGTGGTGGCCAGAACCACTGCGACACCG 2855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTCCTGCGGCGGGAGCTTGGTAACCCTCCAGGTGGTGGCCAGAGCCACTGCGACACCCG 300

Qy   2856 CGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCGACGGAAGTTCCTGAGGAAGAGCA 2915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGCACTGGAAGGCCGCTCCGAAGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCA 360

Qy   2916 GGCGCACCTCGACGCTGATGATTCCAAGGACGTCGCCAAACAGCCTCGTCTGCTGTGTT 2975
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGCGCACTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCGTCTCAACCGGCTGCTGTT 420

Qy   2976 CCCGAAGCCAACCGAAGAGTTCCTCAGGACTGGTCGCAGCACCGTCGGCGAGACTTTGAT 3035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCCGAAGCCAACCGAACGAAGAGTTCCTCGAGCACCGTCGGCGAGACTTTGATCCGCCTGCC 480

Qy   3036 GGATGATCGTGAATTCTTCTACGGACTGGTCACGGCGCTGCGAGGGCGAGACTTTGATCCGCCTGCC 3095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGATGATCGTGAATTCTTCTACGGCTGGTCGTCGAAGGCCGCGAGACTTTGATCCGCCTGCC 540

Qy   3096 AGATGTGGCACCCACTGCTTGTTCGCCTGGATGGGCGATCTCTGAGCCGAGCGATAAGGG 3155
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGATGTGGCACCCCACTGCTTGTTCGCCTGGATGGGCGATCTCTGAGCCGAGCGATAAGGG 600

Qy   3156 TATGCGCAATGTTGTGGCCAAGCTCAACGGGCCAGATCCGCCCAATGCGTGCGTGACCG 3215
          |||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    601 TATGCGCAATGTTGTGGCCAAGCTCAACGGCGTCAACGGGCCAGATCCGCCCAATGCGTGCGTGACCG 660

Qy   3216 CTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCAACAAGGCCCATGT 3275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTCCGTTGAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCAACAAGGCCCATGT 720

Qy   3276 TGCTGCACCATTCGCTGGTGTTGTCACTGACGTGTTGCTGAAGGTGATGAGGTCAAGGC 3335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGCTGCACCATTCGCTGGTGTTGTCACCGGACGTGTTGCTGAAGGTGATGAGGTCAAGGC 780

Qy   3336 TGGAGATGCAGTCGCAATCATCGAGCTATGAGAGATGGAAGCAACAATCACTGCTTCTGT 3395
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TGGAGATGCAGTCGCAATCATCGAGCTATGAGAGATGAAGCAACAATCACTGCTTCTGT 840

Qy   3396 TGACGGCAAGATTGAACCGGTTCCTGCTGCTGCAACGAAGGTGGCAAGGTGGCGACTT 3455
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGACGGCAAAATCGATCGCGGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTT 900

Qy   3456 GATCGTCGTCGTTTCCTAA 3474
          |||||||||||||||||||
Db    901 GATCGTCGTCGTTTCCTAA 919
```

RESULT 14
ABK63659
ID   ABK63659 standard; cDNA; 3945 BP.

```
Query Match        19.6%;  Score 681;  DB 24;  Length 3945;
Best Local Similarity   54.2%;  Pred. No. 1.6e-175;
Matches 1521;  Conservative   0;  Mismatches 1260;  Indels  24;  Gaps   6;

Qy    86 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCAC 145
            ||| | ||   ||  ||  ||| | ||| |  || ||  || ||  || |||| |||
Db   144 TCAAGAAGTAATAGGTGGCCAACAGAGGTGAGATTGCCATCCGAGTGTTTCGTGCCTGCA 203

Qy   146 TCGAAACCGGTGCAGCCACGTAGCTATTACCCCGTGAAGATCGGGGATCATTCCACC 205
            ||  | |  |  ||  ||| | ||| | ||   | || || | |  || ||| ||
Db   204 CAGAGCTGGGTATCCGCACAGTGGCTGTCTACTGGAGACAGAGGCCAGATGCACC 263

Qy   206 GCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGGCTCACCAGTCAAGGCGTACC 265
            || |   ||| | ||||| |  | ||| || |||   || |   |||  ||| | ||||
Db   264 GGCAGAAAGCTGATGAAGCCTACCTTATTGGCCGTGGGCTGGCTCCTGTGCAAGCCTACC 323

Qy   266 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAGTAAAGCAGATGCTATTTACCCGG 325
            || ||| |  ||| | |||  | || || || | || | ||| | || ||| ||||| |
Db   324 TGCACATTCCAGACATCATTAAGGTGGCCAAGGAGAATGGTGTAGATGCTGTGCACCCTG 383

Qy   326 GATATCGGCTTCCGTGCTGAAAATGCCCAGCTTGCCCGCGAGTGCGCGGAAAACGGCATTA 385
            | |  ||| ||||  | || | || |  |   || || |  || | | |  ||  ||| |
Db   384 GCTATGGGTTCCTCTCAGAGAGACAGACTTTGCCCAGGCCTGCTGATAAGTCTCGTGCGGTCC 443

Qy   386 CTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCCGTGCGGTAA 445
            | ||| ||||  ||| ||||||||  |    |  || || || |   || | || ||| |
Db   444 GATTCATTGGTCCAAGCCCAGAGGTCGTCCGCAAGATGGGAGACAAGGTGGAAGCCCGGGG 503

Qy   446 CCGCCGGAAGAAGGTGGTCTGCCAGTTTGGCGGAATCCA---CCCGGAGCAAAAACA 502
            || | |   || | ||  |   | || |||  | |  |||   |||   || |  | ||
Db   504 CCATTGCCATTGCTGCAGGCGGTCCGTGCTGGTCCCTGCCACTAATTCCCCATCAATTCCC 563

Qy   503 TCGATGACATCGTTAAAAGGCTGAAGGCCAGACTTACCCCATCTTGTAAAGGCCAGTTG 562
            || ||| |  |    |  |||     |  || |  ||  | ||||| | |  |||| | |
Db   564 TGCATGAGGCACACGAGTTCTCTAACACCTATGGTTTCCTATTATCTTCAGGCTGCCT 623

Qy   563 CCGGTGGTGGCGGGAGACGGGGTATGCCGGTGCTGGCATGAGGGTGTGCATAGCTACGGAGG 622
            |  ||| ||  | ||| | | |  ||  |  | |||| ||||   || |||| || || |
Db   624 ATGGGAGGTGGGGCGTGGGGAGGCCGGCATGAGAGCTGAAGCGTGAAGCGGCATTCGGAGCTTCGGTATATGTCGAAC 683

Qy   623 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCGAAC 682
            ||  | ||| ||| | |||  || |||| || |  || |  |||| |  || ||||||| |
Db   684 ACACCCGGGCCTTACCCTGAGGCCTTTGGCAATGGGGCATTGTTGTGAGA 743

Qy   683 GTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCCACTGAGAAAG 742
            ||||| | ||| |  || ||  | || |   |||| | ||| || | || || ||||||
Db   744 AATTCATTGAGAGTCGAAGCAAGACACATTGAGGTGCAGATCGTGAGGTGGACCAAATATGGGAACA 803

Qy   743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTGAAA 802
            |||  | || |  ||| |  || || ||| || | ||  | | ||  |  |||||| ||
Db   804 TCTTGCACTTGTATGAGCGGGACTGCTCCATCCAGCGGCGGCCACCAGAAGTGGTAGAGA 863

Qy   803 TTGCGCCAGCACAGCATTGGATCCAGAACTGCGTGATCGCGTGATCGCCGATGCAGTAA 862
            |||| | || ||  | ||| ||| ||| ||| | || ||  |  ||| || |||  |||
Db   864 TTGCCCCTGTCACCCTGACCCCAACTTCGGTCACGCCTCACCAGTGACTCTGTCA 923

Qy   863 AGTTCTGCGCCTCCATTGGTTACCAGGGCGCGGAACCGTGGAATTCTTGGTCGATGAAAG 922
            || |||||| ||||  |||| | || || | ||| || || ||  |||| ||| || ||
Db   924 AACTTGCGGCAGGGTGGCCATGGAGAATGCAGGCACTGTGGAGTTCCTGGTGGACAAGC 983

Qy   923 AGGGCAACCACGTTTTCATGAAGTGAACCCACGTATCCAGGTTGAGCACACCGTGACTG 982
            | ||| || |||  | || || | || | |  | || || || ||| |  | | ||| |
Db   984 ATGGCAAGCACTACTTCATCGAGTCAATTCCGCCGCTGAATTCCAGGTGGAGCACACGGTCACTG 1043

Qy   983 AAGAAGTCACCGAGGTGACCAATGGTGAAGGCGCAGAGCCGCTTGGCTGGTGGCAACCT 1042
            | || || || || || ||| | | | || | |  | || |  || | | | || ||
Db  1044 AGGAGATTACAGATGTGGACCTGGTCCATCCTCAGATCCTCAGATCAAGACCCAGCAGTGCC 1103

Qy  1043 TGAAGGAATTGGCTCGACCCAAGATAAGATCAAGAAAACATCCGAATCAATCAGTTGTC 1102
            || ||| ||||| |  || || |  |  || ||| | || || ||| ||| ||| ||||
Db  1104 TGCCTGACCTAGGCCTTGCGCGGCGGGGAGAGAGAGATCAAGACTGTTGTGCATCAGTTCC 1163
```

XX
AC   ABK63659;
XX
DT   18-JUN-2002 (first entry)
XX
DE   Rat sequence differentially expressed in response to a hepatotoxin #1566.
XX
KW   Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW   differential expression; centrilobular necrosis; steatosis.
XX
OS   Rattus norvegicus.
XX
PN   WO200210453-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US23872.
XX
PR   31-JUL-2000; 2000US-222040P.
PR   02-NOV-2000; 2000US-244880P.
PR   11-MAY-2001; 2001US-290029P.
PR   15-MAY-2001; 2001US-290645P.
PR   22-MAY-2001; 2001US-292336P.
PR   06-JUN-2001; 2001US-295798P.
PR   13-JUN-2001; 2001US-297457P.
PR   19-JUN-2001; 2001US-298884P.
PR   09-JUL-2001; 2001US-303459P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
DR   WPI; 2002-241625/29.
XX
PT   Predicting toxic effects of compounds or the progression of these toxic
PT   effects by determining the changes in gene expression in tissues or
PT   cells exposed to the toxin and comparing these to gene expression in
PT   unexposed tissues or cells -
XX
PS   Claim 1; Seq ID No 1566; 239pp; English.
XX
CC   The invention relates to methods for predicting toxic effects of
CC   compounds or the progression of these toxic effects by determining the
CC   global changes in gene expression in tissues or cells exposed to the
CC   toxin and comparing these to gene expression in unexposed tissues or
CC   cells. Also included are methods of predicting at least one toxic
CC   effect of a compound or progression of a toxic effect, preferably the
CC   hepatotoxicity of a compound, comprising detecting the level of
CC   expression in a tissue or cell sample exposed to the compound of two or
CC   more genes listed in the specification, where differential expression of
CC   the genes is indicative of at least one toxic effect or progression.
CC   The method can also be used to identify an agent which modulates the
CC   toxic response and predict cellular pathways that a compound modulates
CC   in a cell. The methods utilise a set of at least two probes (on a solid
CC   support in kit form), where each of the probes comprises a sequence that
CC   specifically hybridises to a gene listed in the specification, a computer
CC   system comprising a database containing information identifying the
CC   expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC   set of genes comprising at least two genes listed in the specification,
CC   and a user interface to view the information used to present information
CC   identifying the expression level in a tissue or cell of at least one gene
CC   listed in the specification. The method is useful for elucidating global
CC   changes in gene expression and for identifying toxicity markers in
CC   tissues or cell exposed to a known toxin. The genes may be used as
CC   toxicity markers in drug screening and toxicity assays. The genes and
CC   gene expression information may be used as diagnostic markers for the
CC   prediction or identification of the physiological state of tissue or cell
CC   sample that has been exposed to a compound or agent. Hepatotoxicity
CC   is characterised by centrilobular necrosis and steatosis. The present
CC   sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC   which is differentially expressed in response to a hepatotoxic agent.
XX
SQ   Sequence 3945 BP; 883 A; 1108 C; 1127 G; 827 T; 0 other;

```
Qy   1103  GCATCACCACGGAGGAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT  1162
                 |||||||||| |||||||||| |||||||||   | |||||| | |||||||||||
Db   1164  GGGTCACCACTGAGGACCCTGCACCCAGCTTCCACCCAGACACTGGCCCGCATTGAGGTTT  1223

Qy   1163  ACCGGCTCACCAGGCGGAGCGGTTGGCGTTCGTCTTGACGGTG---CAGCTCAGCTCGGTGGCG  1219
                 || |||||||||||| ||| ||||||| || ||||||||   |||||||||||||
Db   1224  TCCGGAGTGGTGAGGGCATGGGCATCCGCTCGGACAATGCCTCAGCATTCCAGGGAGCTG  1283

Qy   1220  AAATCACCGGCACACTTTGACTCATGCTGTGTGAAAATGACCTGCCGTGTTCCGACTTTG  1279
                 || |||||||||||| ||  |||||||||| |||| | |||||||||||||||||
Db   1284  TCATATCCCCCACTATGACTCCCTGCTCGTCTCAAGTCATTGCCCATGGCAAAGACCACC  1343

Qy   1280  AAACTGCTGTTGCTCTGCTGCACAGCGCGGTTGGCTGAGTTCACCGTGTCGTGTTGCAA  1339
                 |||||          |||||||||||||   || ||||||||| |||||||||
Db   1344  CTACAGCTGCCACCAAGATGCAGCAGCCCTGGCGGAGTTCCGTGTCCGAGGGTGTAAAGA  1403

Qy   1340  CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGGAAGAGGACTTCACTTGCCAAGCGCATCG  1399
                 ||||||||||| ||  | | |||||||||||||||| |||||| |||| |||||||
Db   1404  CCAACATCCCCTTCCTGCGAGAATGTGCTCAACACCAGCAGTTCCTAGCGGGCATTGTGG  1463

Qy   1400  CACCGGATTTATCGGCGATCACCACACCTCCTTCAGGCTCCACCTGCGGATGATGAGC  1459
                 |||||| || |||||||| || |||||| ||||| || |||||||| |||||||||
Db   1464  ACACCCAGTTCATCGATGAGACCCCGAGCTGTTCCAGCTGCGCGGCCTGCACGAACCGGG  1523

Qy   1460  AGGGACGCATCCTGGATTACTTGGACAGATGTCACGTGAACAAGCCTCATGTGTGCGTC  1519
                 |||| ||||| |||||||| |||||||| |||||| |||||||||||||||| | |
Db   1524  CCCAGAAGTTGCTACATTACCTTGGACACGTCATGTGTCAATGGCCTACCACTCCAATCC  1583

Qy   1520  CAAAGGATGTTGCAGCACCAATCGATAAGC------TGCCCAACATCAAGGATCTGCCAC  1573
                 |||| ||||   |||||     |||||    |||||||||||||| ||||||||||
Db   1584  CCGTCAAGGTCAGTCCCAGCCCTGGACCCCATTGTTCCTGGTGCCCATTAGGCCCAC  1643

Qy   1574  TGCCACGCGGTTCCCGTGACCGGCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCTGTGATC  1633
                 ||||| || ||| | |||||||||||||| | ||||||| |||||| |||| |||||
Db   1644  CCCCAGCTGGTTCAGAGACATCCTTCTGCGAGAGGGCTTTGCCAGAGCTG  1703

Qy   1634  TCCGTGAGCAGGAGCGACCACTGGCAGTTACTGATCAGCACCAGTCTT  1693
                 || |||||||||||| ||| ||||| ||| ||| |||||||||||||
Db   1704  TGGGGAATCACCAGGGGCCTGCTGCTAATGACACAACCTTCCGGATGCAC  1763

Qy   1694  TGCTTGTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAAAGC  1753
                 || |||| || |||||||| |||| || |||||||| |||||| ||||
Db   1764  TACTTGCCACTAGAGTGCGCACACAGCGATCTCAAAAAGATTGACGTTGCCCACA  1823

Qy   1754  TGACTCCTGAGCTTTTTGTCCGTGGGATCACGAGATGTGGCGATGC  1813
                 ||||||||||||||     |||||  ||||||||||| |||||| |
Db   1884  GCTTCTTGTATGAGTGCCCCTGGCCCCGGCGGCCGCTCCAGGAGCTCCGGGAGCTCCATCCCCAACA  1943

Qy   1874  TAAACATTCAGATGCTGCTTCGCGGCGCCAACACCGTGGGATACACCCCGTACCCAGACT  1933
                 ||||||| ||| || ||| ||| |||||| ||||||| |||| |||||| |||||||||
Db   1944  TCCCATTCCAGATGCTACTGAGGGGGCCAATGCGTGGGCTACACACCTACCCTGACA  2003

Qy   1934  CCGTCTGCGCCGGCGTTTGTTAAGGAAGCTGCCAGCTTCCCGGCTGACATCTTCCGCATCT  1993
                 || ||||||| || ||||||||||||| |||||||||| || |||||||||||||||||
Db   2004  ACGTGGTCTTCAAGTTCTTCTGTGAGTGTGGCCAAAGAGAATGGCATGGACGCTTTCCGGATCT  2063

Qy   1994  TCGACGCGCTTAAACGAGTCTCCAGATGCGTCCAGCAATCGACCGAGACCA  2053
                 ||| ||| ||||||| |  |||||| |  |||| |||| || ||||||||
Db   2054  TTGACTCCCTTAACTACCTGCCAAACATGCTGCTGGGCCATGCTGGGAAGTA  2120

Qy   2054  ACACCCGCGTAGCCGAGGTGGCTATGGCTTATTCTGTGATCTCTGATCCAAATGAAA  2113
                 |||||||| |  ||| ||| || ||| |||  |||| |||||||||||||||||||
Db   2121  CTGGGGGTGTGTGGTGTAGCGTCCATCTACTACACGCGGTCGACGCTGACCCCAGTCGCA  2180

Qy   2114  AGTCTACACCCTGATTACTACCTAAAGATGGCAGAGAGATCGTCAAGTCTGGCGCTC  2173
                 ||| |||| |||||| |||| |||||| | ||  |||  || || || |||| ||||
Db   2181  CTAAATACTCACTGGAGTACTACATCCACGTTAGCTGAAGAACTGGTGCGAGCCCGGCACTC  2240

Qy   2174  ACATTCTGGCCATTAAGGATATGGGCTGGTCTGCTTCGCCCAGCTGCGCGGTAACCAAGCTGG  2233
```

```
Db   2241  ACATCCTCTGCATTAAGGACATGCGAGCCTGCTGCTGGAAGCCTGCGAAGCCTGGCACCATGCTGG  2300
                 |||| || |   |||||||| ||| | ||||  |||||| |||||   ||||||||||||
Qy   2234  TCACCGCACTGCG----CCGTGAATTCGATCTGCCAGTGCACGTGCACCCACCCACGACACTG  2290
                 || | || ||    |||| | |||| || ||||  |||||| ||  ||| | ||||||
Db   2301  TCAGCTCCCTCCGGACCGGTTCCCGACCCTCCACTGCACATCCATCCATGACACAT  2360

Qy   2291  CGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAGCTGGTGTGCAGATGCTGTTGACG  2350
                 | |  ||| | ||| ||||| ||| ||| | |||||| | | ||| ||||| |||| |||
Db   2361  CAGGGTCAGGTGTGGCAGCCACATGTTGGGCTGTGCAGACTGGGCTGTGATGTGTGGGATG  2420

Qy   2351  GTGCTTCCGCACCACTGTCTCGGCACCACCTCCCAGCCATCCCGTCTGCCATTGTTGCTG  2410
                 ||||||| ||||| || |||||||||||| ||| ||||||||  |||||| | |||| |
Db   2421  TGGCAGTCGACTCTATGTCTGGGATGACCTCACAGCCCTGGGGGCCCGTGGGCCT  2480

Qy   2411  CATTCGGCGGCACACCCGTGCGGCGATACCGGTTTGGCCTCGTTTCTGACCTGAGC  2470
                 || ||| ||||| |||| |  | | |||| |  ||||||||| || |||||||||
Db   2481  GTACCAAAGGGACTCCTCTGGACACAGAGGTACCCCTGGAGCGTGTGTTTGACTACAGTG  2540

Qy   2471  CGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTGA------GTCTGGAACCCCCAG  2524
                 ||||||||||| |  | || ||||| | |||||||| |      ||||||||| ||||
Db   2541  AGTATTGGGAAGGGGCTGGGGGGCTGTATGCGGCCCTTTGATTGCACGGCTACCATGAGT  2600

Qy   2525  GCCCAACCGGTCCGGTCTCACCGGCCACGAAATCCAGGCGGACAGTTGTCCAACCTGCCTG  2584
                 |||||  |||| ||||||| || | | |  |||| ||| | || || |||||| |||||
Db   2601  CTGCAACTCAGACGTGTATGAGAATGAGGATCCAGGGGGCCCAGTACACCAACCTACACT  2660

Qy   2585  CACAGGCCACCGCACTGGGCGTTGCTGCTCATCGAGACACTACCCAG  2644
                 |||| ||| || |||||| |||||| || || |  |||||||| |||
Db   2661  TCCAGGCCCACAGCAATGGGACTTGGCTCCAAGTTCAAGGAGGTCAAGAGGGCTATGTGG  2720

Qy   2645  CCGTTAATGAGATGCTGGGACAGCGCCACCAAGGTCACCCATCCTCCAAGGTTGTTGGGCG  2704
                 ||||||||||| | |||| | |  |||||||||| |||||||||||||||||||| |||||
Db   2721  AGGCTAACCAGATGCTGGGGGACCTCATCAAGGTGACCAACATCCTCCAAGATTGTGGGGG  2780

Qy   2705  ACCTCGACTCCACCTGGTTGGTGCGGGTGTAGATCCAGACAGACTTTGCTGCGAGACCCAC  2764
                 |  |||  |||||||| ||| || | | |  ||| ||| |||||||||||| |||||||||
Db   2781  ATCTGGCCCAGTTCATGTGCAGAACGGGTTGAGCCGGGAGAACGGGCAGCAGAAGTCAGGCAG  2840

Qy   2765  AAAAGTACGACACATCCCAGACTCTGTCATCCGGTTCGTCTGCGGCGCGGCGCGAGCTTGGTAACCCTC  2824
                 |  |  | |||| |||  ||||  |||| | |   ||  |||||||| | |  ||||  || |
Db   2841  AAGAGCTGTCCTTCCCCCCGCTCTGTGGTGGAGTTCCTGCAGGGCTACATATTGGCATTCCCC  2900

Qy   2825  CAGGTGGCTGGCTGGCCAGGAACCACTGCGCCACCCGGCGCACTGGAAGGCC  2869

Db   2901  ATGGGGGTTTCCCTGGAACCCTTCCGTTCTAAGGTGCTAAAGGACC  2945
```

```
RESULT 15
ABN96532
ID   ABN96532 standard; DNA; 4017 BP.
XX
AC   ABN96532;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #3030 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
```

XX
PI   Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer,
PT   hepatocellular carcinoma or metastatic liver tumor in a patient,
PT   involves detecting the level of expression of two or more genes in a
PT   liver tissue sample -
XX
PS   Claim 1; SEQ ID NO 3030; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 4017 BP; 825 A; 1238 C; 1235 G; 719 T; 0 other;

Query Match   19.4%;   Score 674.2;   DB 24;   Length 4017;
Best Local Similarity   54.33%;   Pred. No. 1.2e-173;
Matches 1523;   Conservative   0;   Mismatches 1258;   Indels   24;   Gaps   7;

Qy   86 TCAAAAAGATCTTGGTAGCAAACCGCCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCAC 145
          || || || ||||| || || |||| | ||| ||| |||| || ||| || ||||
Db   192 TCAAGAAAGTCATGGTGGCCAACAGAGTGGAGATTGCCATCCGTGTTCCGGGCCTGCA 251

Qy   146 TCGAAACCGGTGCAGCCACCGTAGCTATTTACCCCGTGAGAATCGGGGATCATTCCACC 205
          | ||| ||||| |||| ||||| |||| || | ||| ||| || |||| | |||
Db   252 CGGAGCTGGGGGCATCCCACCGTAGCCATCTACTCTGAGCAGGACACACGGGATGCACC 311

Qy   206 GCTCTTTTGCTTCTGAAGCTGTCCGCCATTGGTACTGAAGGCTCACCAGTCAAGGGGTACC 265
          |||| | || | |||| || || || || ||||||||||| |||||||| || ||||||
Db   312 GGCAGAAAGCAGATGAAGCCTATCTCATCGGCCGCGGACCTTCCGCCCCGTGCAGGCCTACC 371

Qy   266 TGGACATCGATGAAGAAATTATCGGTGCAGCTAAAAAAGTGAAGCAGATGCTATTTACCCGG 325
          || ||||| || || || || ||||| || || |||| |||| | || | | || |||
Db   372 TGCACATCCCAGACATCATCAAGGTGGCCAAGGAGAACAACGTAGCAGTGCCACCCTG 431

Qy   326 GATATGGCTTCCTGTCTGAAAATGCCGAGTTGCCGGCGAGTGCGCGGAAAACGGCATTA 385
          | | || |||| | ||||| | ||| || | |||| | |||| || || || | | |
Db   432 GCTACGGGTTCCTCTCTGAGCGAGGGGACTTCGCCGGAGGCCGGCAGGGTCC 491

Qy   386 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTAA 445
          | ||||||||| |||| ||||| ||| || |||| | ||||| || | | || || |
Db   492 GGTTTATTGGTCCAAGCCCAGAGTTGTCCGGAGGACAAGGTGAGGCCCGGG 551

Qy   446 CGGCCGGAAGAAGGCTGGGTGCCCAGTTTTTGGCCG---AATCCACCCCGAGACAAAAACA 502
          | ||||| || || | |||| ||||||| ||| | | | || |||| || | |
Db   552 CCATCGCCATTGCCGGGGGTGTTCCGGTCCTTGCACAGATGCCCCATCACGTCC 611

Qy   503 TCGATGACATCGTTAAAGGCCTGAAGGCCGAGACTTACCCCATCTTTGTAAAGGCAGTTG 562
          | ||||| |||| | ||| || ||| ||||||| |||||||| | | ||| |||| |
Db   612 TGCATGAGGCCCACAGGTTCTCCAACACCTACGGCTTCCCCATCATCTTCAAGGCGGCCT 671

Qy   563 CCGGTGGTGGCGGACGGCGTATGCCGTTTGTTCTTCACCTGATGAGCTCCGCAAATTGG 622
          || ||||| || ||| | | || | | || |||| | | ||| | ||| || |
Db   672 ATGGGGGGTGGGGGGCTGGCATGAGGGGTGCACAGGCTGGTAGCAGGCTGAGAGGAGAATT 731

Qy   623 CAACAGAAGACATCTCGTGAAGCTGGCCATTCGGCGACGGTTCGGTATATGTCGAAC 682
          | || || || | ||| | || | |||| | | | ||| | ||| | | ||| ||
Db   732 ACACCCCGGGCCTACTCAGAGCCCTGCCGCGGGGCCTTTGGGAATGGGGGCTGTTTGTGGAGA 791

Qy   683 GTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCCACTGGGAGAAG 742
          | || |||| ||| ||||| |||| || ||||| | ||| |||| ||| |||||||
Db   792 AGTTCATCGAGAAGCCACGGCACCGCAGGTGCAGATCTTGGGGGACCAGTATGGGAACA 851

Qy   743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCCTCGTCACTGCAGCCTCGTCGAAA 802
          || || || || | |||| ||| |||| | | || ||| || |||| | ||
Db   852 TCCTGCACCTGTACGCTACGGAGGCGAGACGTGTCCATCCAGCCACCGGCCACCAGAGGC 911

Qy   803 TTGGCGCAGCACAGCATTTGATCCAGAACTGCGTGATCGGCGACTCGGCTCACCAGGACTCTGA 862
          || ||| | | || | | | |||| || |||| || ||| | ||| || | |||| ||
Db   912 TTGCCCCGCCGCCCACCTGGACCCGGCTTCGGCGTCAGCCAGTTGTGCCACCGACTCTGA 971

Qy   863 AGTTCTGCCGCTCCATTGGTCCGCCGGCGGCGGGAACCTGGAATTCTTGGTCTGATGAAA 922
          || | ||| | || | || || | | | ||| || ||| | | ||| | || | |||
Db   972 AACTCGCTAAACAGGTGGGCTACGACAACGAGGCACCGTGGAGTTCCTGGTGGACACGGGC 1031

Qy   923 AGGGCAACCACGTTTCATCGAAATGAACCACCAGTATCCAGGTTGAGCACGTGACTG 982
          || || || ||| | |||| || ||| |||| ||| ||||| |||| | ||| ||
Db 1032 ACGCAAGCACTACTTCATCGAGGTCAACTCCCGCCTGCAGTGGGAGCACACACGGTCACAG 1091

Qy   983 AAGAGTCACCGAGGTGACCTGGTGAAGGCGCAGATGCGCTTGGCTGTGCTGCAACCT 1042
          || | ||||| | || || |||||| | || || | | || | |||| || |||
Db 1092 AGGAGATCACCGACGTGGTCCATGGTCAGATAGATGAGATCAGATCTCAGATCCACGGAGGCC 1151

Qy 1043 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACAATCCGCATCAACGGGTGTGCATCCAGTGCC 1102
          |||| | | || | | | ||||| |||||||| | | | || | || | || |||| || |||
Db 1152 TACCCGACCTGGGCTGCGCGGAACATCCGCATCAACGGGTGTGCCATCCAGTGCC 1211

Qy 1103 GCATCACCACGGAAGATCAAACACAACGGCTTCGCCCAGATACCGGAACTATCACCCGT 1162
          || || || | || |||| || || | || | || | | |||| | |||||
Db 1212 GGGTCACCACCGAGGACACCCCGGCGTTCGTGTCTGA---TGACGGTGCAGTCGAGTGCATTGAGGTGT 1271

Qy 1163 ACCGCTCACCAGGCGGAGCTGGCGTTCGTGT---TGACGGTGCAGTCGAGTCTCGGTGGCG 1219
          || || || | || |||| ||| || |||| | | || | | ||| | | | || ||| |
Db 1272 TCCGGAGCGGAGGGCATGGGCATCCGATAATGCTTCCGGCCTTCCGTGGTTCCGGGCCTGCA 1331

Qy 1220 AAATCACCGCACACTTTGACTCCATGTGGTGAAAATGACCTGCTGAGTTGAAAATGACCACC 1279
          || | || | || || |||| | | || | | |||| || | || | |||| ||||
Db 1332 TCATCTCGCCCCCACTACGACTCCTGGCCTGCTGTCTCAGTGTCACCGGCCCTTGCAGTGCACC 1391

Qy 1280 AAACTGTGTTGCTCGTGCACAGCGCGCCGTTGCTCGAGTTCACCGTGTTGTTGCAA 1339
          ||| |||||| | ||| ||| || || | | | |||| ||| || | ||||
Db 1392 CCACGGCCGCGCCCTTGCCCAAGATGAGCAGGGCCCTGTGCCGTCCGAGTGTGAAGA 1451

Qy 1340 CCAACATTGGTTTCTTGCGTGCTCTGCAGATATGCTCAGCGTCCAAGGACCATCG 1399
          ||||||| | || | |||| | |||| |||| || || ||| ||| |||||
Db 1452 CCAACATCGCCTTCCTGCGTGCTCAACAACCAGCAGTTCCTGCAGCCACTGTGG 1511

Qy 1400 CCACCGGATTTATCGCCAGCACCCGATCACCCACACCTTCAGGCTTGCGGATGATGAGC 1459
          || || | | |||| ||| | || |||| ||| || | || || || | ||||||||
Db 1512 ACACCCAGTTCATCGACGAGAACCCAGACGCGTGTTCCAGCTGCGGCCTGCGCACAGAACCGGG 1571

Qy 1460 AGGGACGACATCCTGGATTACTTGGCAGATGTCGAA---CAAGCCTCATGGTGTG 1515
          || || | ||||| || ||| | | |||| ||| | |||| || | ||||| |
Db 1572 CCCAAAAGCTGTTGCACTACTTGGCCCATGTCCATGGAAGTAAACGGTCACCACCCCGATTC 1631

Qy 1516 CGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATCAAGGATCTGCCACTG 1575
          || |||| || | || | | | |||| | || |||| | |||||| | | |||||||
Db 1632 CCGTCAAGGGCCAGCCCGTTGTTCCTGACGACCCCGTTGTCGTGTCATGTAGGGGTCGGG 1691

Qy 1576 CCACG--CGGTTCCCGTGACCGCCTGAAGCAGCGCGCGTTTTTGTCTGTATC 1633
          ||||| ||| | | | | || ||||||| || | | | | || | |||
Db 1692 CCCCGGTCGTTTTCAGAGACATCCTGCTCGCCAGCATCAAACCGTCCCATAGGCCCGG 1751

Qy 1634 TCCGTGAGCAGGACGCACTGGCCAGTTACTGACACCTTCCGCGATGCAGTTCTTCGAGCTG 1693
          |||| || || | || |||| | ||||| | || || ||| | || | | | | ||
Db 1752 TGCGGAACACCCGGGCTGCTGTGACGACCACCAGTTCAGGGACGACCAGACCTTCAGGAGCCCACCAGTCAC 1811

Qy 1694 TGCTTGGCGACCCGAGTCCGCTCATTCCGCACTCAGGATATATGTGAAC 1753
          ||| | || ||| ||| || | | |||| | | | ||| | |||| |
Db 1812 TGCTGGCCACTCGTGTGCGCACCACCGACGAATCTCAAAAAGATCGCCCCCCATGTTGCCACA 1871

us-09-974-973-1.rng

Wed Apr 9 09:38:21 2003

```
Db  2949 ATGGGGGTTTCCCGAACCCTTTCGCTCTAAGGTACTGAAGGACC 2993
              ||||| || ||  || |||| |||| ||    ||||||||||||
```

Search completed: April 5, 2003, 05:53:14
Job time : 1143 secs

```
Qy  1754 TGACTCCTGAGCTTTTGTCCGTGGAGGCCTTGGGCGCGGCGACCTACGATGTGGCGATGC 1813
              ||||||||||||    |  ||||||| || |||  |||||| |||| |||||| |||||
Db  1872 ACTTCAGCAAGCTCTTCAGCATGGGAGAACTGGGAGGAGCCACGTTTGACGTCGCCATGC 1931

Qy  1814 GTTTCTCTTGAGGATCCGTGGGACAGGCCGGCGGCTGCGGAGGCGGATGCCGAATG 1873
              | ||| |  ||||   |  |||| ||  |||  |||||| ||||| ||| |||||
Db  1932 GCTTCCTGTATGAGTGCCCCTGGCGCGGCTGCGAGAGCTCCGGGAGCTCATCCCAACA 1991

Qy  1874 TAAACATTCAGATGTGCTTCGCCGGCCAACACCGTGGGATACACCCGTACCCAGACT 1933
              |  | || |  | ||| ||  |   |||| ||||||| |  ||| |||| |||
Db  1992 TCCGTTTCAGATGCTGCTGCGGGGGGCCAATGCTGTGGCTACACCAACTACCCAGACA 2051

Qy  1934 CCGTCTGCCGCCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGCATCT 1993
              || ||||||||| |  |  | | |||| | |   |  ||| |||    |  || ||| |
Db  2052 ACGTGGTCTTCAAGTTCTGTGAAGTGGCCAAAGAGAATGCATGGATGTCTTCCGTGTGT 2111

Qy  1994 TCGACGGCGCTAACGACGTCTCCCAGATGCGTCAGCAATGCAGCGCAGTCCTGGAGACCA 2053
              | |||| | || ||| |  |  ||| ||  |  ||||  |  |   ||| |  ||||||
Db  2112 TTGACTCCCTCAACTACTTGCCCACAATGCTGCTGGGCATGACGTGCTGGGCG---CAGGAAGTG 2168

Qy  2054 ACACCGCGGTAGCCGAGGTGCCTATGGCTTATTCTGTGATCTCTCTGATCAAAATGAAA 2113
              |  ||| ||| | | ||| | |  |    |  |  ||   ||  |  ||   |||||
Db  2169 CCGGAGGCGGTGTGGACAGGCTGCCATTCATACACGGGCGACGTGGCCGCGCTGGCACCCA 2228

Qy  2114 AGTCTCACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTCGGCGCTC 2173
              | |  ||||| || ||  |  |    |   ||  |  ||| || |  |||| ||| ||| |
Db  2229 CCAAGTACTCACTGCAGTATACACATGGGCTTGGCTGGCTGCGGCAGCTGGCACCC 2288

Qy  2174 ACATTCTGGCCATTAAGGATATGCGTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGG 2233
              |||| | ||||  |  ||| |||| ||||||| || ||||||||  || |||| |||||
Db  2289 ACATCCGTGCATCAAGGACATGGCCCGGGCTGCTGAAGCCCACAGGCCTGCTGAACCATGCTGG 2348

Qy  2234 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTG 2290
              ||  |||| |||    ||  |  || |||  ||||| | || || |  ||  ||| |||
Db  2349 TCAGCTCCCTGGGACACCGTCCCCGACCTCCCACTGCACATCCACGCACACCGT 2408

Qy  2291 CGGGTTGGCCAGTTGGCTACCTACTTTGCTGCCTGGCTCCAGCTCAAGCTGGTGCAGATGCTGTTGACG 2350
              ||||  || ||||  |   |     ||| ||||  |  ||| | |||||| || ||| || ||  ||
Db  2409 CAGGGGCAGGGGTGGCAGCCGGTCGGGGACTGGGCAGCTGGCTGTGTCGTGTGTGGATG 2468

Qy  2351 GTGCTTCCGGACACTGTCTGCGCACCACTCCCAGCCATCCTGTCTGCCATTGTTGCTG 2410
              ||||||||  ||||| |  | || ||  ||  || |||   || || ||| || |||
Db  2469 TGGCAGCTGATTCCATGTCTGGGATGACTTCACAGACCCTGGCCCCTGGGTGCGCCT 2528

Qy  2411 CATTCCGCGCACACCCCTGGCGGATACCGGTTTGAGCCTCGAGGCGTTGTTTCTGACCTCGAGC 2470
              || ||| | ||| ||| | ||  ||| |  ||||  ||||| ||  | || | ||| |||||
Db  2529 GTACCAGAGGGACTCCCCTGGACACAGAGTGCGGGGCTGTTTGACTACACAGTG 2588

Qy  2471 CGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGACTTCTGGAACCCCAGGCCCAA 2530
              ||| ||||  | | | |     |||||| ||||   |||   |  || |||||||||||
Db  2589 AGTACTGGGAGGGGCTTGGGCGTCGGGGACTGTACCGGCGCCCTCGACTGCACCATGAAGT 2648

Qy  2531 CCGG------TCGCGTCTACCGCCACGAAATCCCAGGCGCAGTTGTCCAACCTGCGTG 2584
              || ||      | | ||||   |   |  |||| | |  |||  |||||| |||||
Db  2649 CTGGCAACTCGGACGTGTATGAAAAATGAGATCCCCGGCGTCATCGAAGACAACTACGCAG 2708

Qy  2585 CACAGGCCACCGCACTGGGCGCTTGCTTGATCGCTTCATCGCTTCGAGCTCATCGAAGAGGCCTATGTGG 2644
              ||| |||  |||||| |||  ||  ||   |   |   ||   ||  ||| ||| ||||||
Db  2709 TCCAGGCCCACAGCAGCTGGGCTTGGCTGGCTGCTCCCAAGTTCAAGGAGGTCAAGAGAGGCCTATGTGG 2768

Qy  2645 CCGTTAATGAGATGTCTGGGACGCCCAACCAAGTCACCCCATCCTCCAAGGTTGTTGGCG 2704
              || || ||| | |  |||| | |    || |  |   || | | || ||||  |||  |
Db  2769 AGGCCAACCAGATGCTGGGCGCATCTCATCAAGGTGACGCCCTCCTCCAAGATCGTGGGGG 2828

Qy  2705 ACCTCGCACTCCACCTGGTTGGTGCGGGTGGTCGCAGATCTGGGTGCGCCGAGACCCAC 2764
              |||| ||| | ||  ||| ||||| ||  |  |  |  |  |  || ||  | | |||
Db  2829 ACCTGCCCAGTTTATGGTGCAGAATGGATTGAGCCCGGTGATTGAGCTCAGGCGG 2888

Qy  2765 AAAAGTACGACATCCCAGACTCTGTCATCGGCTTGGTAACCCTC 2824
              |   || | |  |  |  |   |  |   ||| |  |||  | |
Db  2889 AAGAGCTGTCCTTTCCCCGCTCCGTGGAGTTCCTGCAGGGCTACATCGGTGTCCCCC 2948

Qy  2825 CAGGTTGGCTGCCAGAACCACTGCCACCCCGCGACCTGGAAGGCC 2869
```

```
                    GenCore version 5.1.3
         Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 5, 2003, 07:14:43 ; Search time 297 Seconds
                                (without alignments)
                              10260.205 Million cell updates/sec

Title:         US-09-974-973-1
Perfect score: 3474
Sequence:      1 gtgactgctatcacccttgg.........tgatcgtcgtgttcctaa 3474

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched:      593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters:      1186858

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published_Applications NA:*
         1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
         7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
         8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
        10: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
        12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
        13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
        14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3474	100.0	3474	9	US-09-974-973-1	Sequence 1, Appli
2	3474	100.0	3474	9	US-09-974-973-3	Sequence 1, Appli
3	3398.8	97.8	3621	9	US-10-045-072-1	Sequence 1, Appli
4	3344.8	96.3	3420	9	US-09-738-626-765	Sequence 765, App
5	681	19.6	3945	10	US-09-917-800A-1566	Sequence 1566, Ap
6	674.2	19.4	4017	10	US-09-880-107-3029	Sequence 3029, Ap
7	502.8	14.5	3429	10	US-09-815-242-6709	Sequence 6709, Ap
c 8	488.4	14.1	5030	7	US-08-781-986A-324	Sequence 324, App
c 9	488.2	14.1	20072	10	US-09-070-927A-89	Sequence 89, Appl
10	486.8	14.0	3441	10	US-09-815-242-4413	Sequence 4413, Ap
11	428.8	12.3	3222	10	US-09-815-242-8263	Sequence 8263, Ap
12	305	8.8	1362	10	US-09-767-479-7	Sequence 7, Appli
13	302.2	8.7	1350	10	US-09-815-242-6233	Sequence 6233, Ap
14	297	8.5	3065	10	US-09-767-479-5	Sequence 5, Appli
15	286	8.2	2238	10	US-09-974-300-2	Sequence 2, Appli
16	283.8	8.2	1350	10	US-09-815-242-9787	Sequence 9787, Ap
17	259.6	7.5	1347	10	US-09-815-242-7063	Sequence 7063, Ap
18	258.6	7.4	1416	10	US-09-815-242-4160	Sequence 4160, Ap
19	255	7.3	1350	10	US-09-815-242-7965	Sequence 7965, Ap
20	249	7.2	1368	10	US-09-815-242-9519	Sequence 9519, Ap
21	245.8	7.1	1368	10	US-09-815-242-9266	Sequence 9266, Ap
22	243.4	7.0	1773	9	US-09-738-626-3440	Sequence 3440, Ap
23	234.2	6.7	1377	10	US-09-815-242-7224	Sequence 7224, Ap
24	231	6.6	1368	10	US-09-815-242-7461	Sequence 7461, App
25	229	6.6	6021	10	US-09-070-927A-458	Sequence 458, App
26	225.8	6.5	1329	10	US-09-815-242-3908	Sequence 3908, Ap
27	225.8	6.5	1371	10	US-09-815-242-6827	Sequence 6827, Ap
28	203.4	5.9	1335	9	US-09-891-641-59	Sequence 59, Appl
29	191.4	5.5	2175	9	US-10-160-501-18	Sequence 18, Appl
30	191.4	5.5	2577	9	US-10-160-501-16	Sequence 16, Appl
c 31	184	5.3	1359	9	US-09-738-626-777	Sequence 777, App
32	180.8	5.2	989	10	US-09-974-300-5516	Sequence 5516, Ap
c 33	179.2	5.2	3710	7	US-08-781-986A-405	Sequence 405, App
34	178.8	5.1	1356	10	US-09-815-242-9029	Sequence 9029, Ap
35	167.2	4.8	1346	10	US-09-815-242-4751	Sequence 4751, Ap
c 36	164.4	4.7	2719	12	US-10-007-693-136	Sequence 136, App
37	163.2	4.7	1338	10	US-09-815-242-4363	Sequence 4363, Ap
38	163.2	4.7	1362	10	US-09-815-242-8464	Sequence 8464, Ap
c 39	162.8	4.7	2785	7	US-08-781-986A-133	Sequence 133, App
40	162	4.7	2205	9	US-09-938-842A-2078	Sequence 2078, Ap
41	158.4	4.6	1330	10	US-09-974-300-1081	Sequence 1081, Ap
42	143.8	4.1	555	10	US-09-974-300-1108	Sequence 1108, Ap
43	140.2	4.0	1077	10	US-09-815-242-8841	Sequence 8841, Ap
44	119	3.4	546	10	US-09-974-300-1106	Sequence 1106, Ap
45	106.2	3.1	474	10	US-09-815-242-3976	Sequence 3976, Ap

ALIGNMENTS

```
RESULT 1
US-09-974-973-1
; Sequence 1, Application US/09974973
; Patent No. US2002017202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteriu
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3474)
US-09-974-973-1

  Query Match             100.0%;  Score 3474;  DB 9;  Length 3474;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  3474;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACTGCTATCACCCTTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy     61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CACACATCTTCAACGCTTCCAGCATTCAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120

Qy    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTACCCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTACCCC 180

Qy    181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
```

```
Qy  1321  ACCGGTCGTCGGTGTTGCAACCACATTGGTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  ACCGGTCGTCGGTGTTGCAACCACATTGGTTCTTGCGTGCGTTGCTGCGGGAAGAGGAC  1380

Qy  1381  TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TTCACTTCCAAGCGCATCGCCACCGGATTTATCGGCGATCACCCACACCTCCTTCAGGCT  1440

Qy  1441  CCACCTGCGGATGATGAGCAGGAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  CCACCTGCGGATGATGAGCAGGAGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAAC  1500

Qy  1501  AAGCCTCATGGTGTCGCGTCCAAAGGATGTTGAGCACCAAATCGATAAGCTGCCCAACATC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  AAGCCTCATGGTGTCGCGTCCAAAGGATGTTGAGCACCAAATCGATAAGCTGCCCAACATC  1560

Qy  1561  AAGGATCTGCCACTGCCACGCCGGTTCCGTGACCGCTGAAGCAGCGCCAGCCCGCG  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  AAGGATCTGCCACTGCCACGCCGGTTCCGTGACCGCTGAAGCAGCGCCAGCCCGCG  1620

Qy  1621  TTTGCTCGTGATCTCGTGAGCAGGAGACCCACTGGCAGTTACTGATACCACCTTCCGCGAT  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  TTTGCTCGTGATCTCGTGAGCAGGAGACCCACTGGCAGTTACTGATACCACCTTCCGCGAT  1680

Qy  1681  GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGACTGAAGCCTGCGCACTGGAGCAGAG  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  GCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGACTGAAGCCTGCGCACTGGAGCAGAG  1740

Qy  1741  GCCGTGCGAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCGCTGGGCGGCGGCGGCGGCGACTTAC  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  GCCGTGCGAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCGCTGGGCGGCGGCGGCGGCGACTTAC  1800

Qy  1801  GATGTGGCGGATCGCGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGGAG  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  GATGTGGCGGATCGCGTTCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGGAG  1860

Qy  1861  GCGATGCCGAATGTAAACATTCAGATGCGTTCGCGGCCGCAACACCGTGGGATACACCC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  GCGATGCCGAATGTAAACATTCAGATGCGTTCGCGGCCGCAACACCGTGGGATACACCC  1920

Qy  1921  CCGTACCCAGACTCCGTCGCCGCGCGTTGTTGAAGCTGCCAGTCCGCGTGGAC  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CCGTACCCAGACTCCGTCGCCGCGCGTTGTTGAAGCTGCCAGTCCGCGTGGAC  1980

Qy  1981  ATCTTCCGCATCTTGACGACGCTTAACGACGCTCCCAGATGCGTCAGCAATCGACGCA  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ATCTTCCGCATCTTGACGACGCTTAACGACGCTCCCAGATGCGTCAGCAATCGACGCA  2040

Qy  2041  GTCCTGGAGACCAACACCGGTAGCCGGTAGCCGCAGATGGCTATTCGGTGATCTCT  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GTCCTGGAGACCAACACCGGTAGCCGGTAGCCGCAGATGGCTATTCGGTGATCTCT  2100

Qy  2101  GATCCAAATGAAAAGCTCTACACCTACTAAAGATGCGAGAGAGAGATCGTC  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  GATCCAAATGAAAAGCTCTACACCTACTAAAGATGCGAGAGAGAGATCGTC  2160

Qy  2161  AAGTCTGGCGCGCTCACATTCTGGCCACTGGCTGTCTGCTTCGCCCAGCTGCG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  AAGTCTGGCGCGCTCACATTCTGGCCACTGGCTGTCTGCTTCGCCCAGCTGCG  2220

Qy  2221  GTAACCAAGTCTGGTCACCGACGCCCATTAGGATATGGCTGTCGCAGTTGCACACC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  GTAACCAAGTCTGGTCACCGACGCCCATTAGGATATGGCTGTCGCAGTTGCACACC  2280

Qy  2281  CACGACACTGCGGGTGGCCAGTTGGCTACCTTTGCTGCAGCTCAAGCTGCAGAT  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CACGACACTGCGGGTGGCCAGTTGGCTACCTTTGCTGCAGCTCAAGCTGCAGAT  2340

Qy  2341  GCTGTTGACGTGCTTCCGGACCACTGTCTGCGGCACCACTGTCTCTCGCCTCCGCC  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCTGTTGACGTGCTTCCGGACCACTGTCTGCGGCACCACTGTCTCTCGCCTCCGCC  2400

Qy  2401  ATTGTTGCTCATTCGCCGATACCGGTTTGAGCCTCGATACCGGTTGACGCTGTTTCT  2460
```

```
Qy   241  GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGTGCAGCTAAAAAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGTGCAGCTAAAAAA  300

Qy   301  GTTAAAGCAGATGTCTATTTACCCGGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GTTAAAGCAGATGTCTATTTACCCGGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC  360

Qy   361  CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCAACCCCAGAGGTTCTTGATCTC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGGCCAACCCCAGAGGTTCTTGATCTC  420

Qy   421  ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGGAAGAGGCTGGTCTGCCCAGTTTTGGCG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  ACCGGTGATAAGTCTCGTGCGGTAACCGCCGCGGAAGAGGCTGGTCTGCCCAGTTTTGGCG  480

Qy   481  GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTAC  540

Qy   541  CCCATCTCTTTGTAAAGGCAGTTGCCGGTGGGCCGGACGCGGTATGCGGCTTTGTTCTTCA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CCCATCTCTTTGTAAAGGCAGTTGCCGGTGGGCCGGACGCGGTATGCGGCTTTGTTCTTCA  600

Qy   601  CCTGATGAGCTCCGCAAATGGCAACAGAAGCATCTCGTGAGCTGAAGCGGCATTCGGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CCTGATGAGCTCCGCAAATGGCAACAGAAGCATCTCGTGAGCTGAAGCGGCATTCGGC  660

Qy   661  GACGGTTCGGTATATGTCGTGATTAACCCCCAGACAATTGAAGTGCAGATC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GACGGTTCGGTATATGTCGTGATTAACCCCCAGACAATTGAAGTGCAGATC  720

Qy   721  CTTGCGATCGGACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTTGCGATCGGACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780

Qy   781  CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT  840

Qy   841  CGCATTGTGCGGATGCAGTAAAGTTCTGCGCGCTCCATTGGTTACCAGGGCGCGGGAACC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCATTGTGCGGATGCAGTAAAGTTCTGCGCGCTCCATTGGTTACCAGGGCGCGGGAACC  900

Qy   901  GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATC  960

Qy   961  CAGGTTGAGCACACCGTGACTGCACCGTGAAGGAAGTCACCGAGGTGGACTTGGTGAAGGCGCAGATG  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CAGGTTGAGCACACCGTGACTGCACCGTGAAGGAAGTCACCGAGGTGGACTTGGTGAAGGCGCAGATG  1020

Qy  1021  CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1080

Qy  1141  GATACCGGAACTATCACCGGCTGCAGCACTGCCAGTGCGCATCACGGCTACCGGTTCGTCTTGACGGT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GATACCGGAACTATCACCGGCTGCAGCACTGCCAGTGCGCATCACGGCTACCGGTTCGTCTTGACGGT  1200

Qy  1201  GCAGCTCAGCTGCCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTCGGTGAAAATGACC  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCAGCTCAGCTGCCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTCGGTGAAAATGACC  1260

Qy  1261  TGCCGGTTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGGTTGGCTGAGTTC  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  TGCCGGTTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCGGTTGGCTGAGTTC  1320
```

US-09-974-973-3
; Sequence 3, Application US/09974973
; Patent No. US2002017202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteriu
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 3
; LENGTH: 3474
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-3

Query Match   100.0%;   Score 3474;   DB 9;   Length 3474;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 3474;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 GTGACTGCTATCACCCTTGGCGGTCTCTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTGACTGCTATCACCCTTGGCGGTCTCTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60

Qy   61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGCGGCGAAATC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGGCGGCGAAATC 120

Qy   121 GCGGTCCGTGCTTCCGTGCAGCACTCGAACCGGTGCAGCGTAGCTATTACCCC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCGGTCCGTGCTTCCGTGCAGCACTCGAACCGGTGCAGCGTAGCTATTACCCC 180

Qy   181 CGTGAAGATCGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CGTGAAGATCGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240

Qy   241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCCGTCGTGCAGCTAAAAAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCCGTCGTGCAGCTAAAAAA 300

Qy   301 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360

Qy   361 CGCGAGTGCGCGGAAAACGGCATTACTTTATTGGCCAACGTCTTGATCTC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CGCGAGTGCGCGGAAAACGGCATTACTTTATTGGCCAACGTCTTGATCTC 420

Qy   421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAGAAGGCTGGTCTGCCAGTTTCGGCG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAGAAGGCTGGTCTGCCAGTTTTGGCG 480

Qy   481 GAATCACCCCGAGCAAAAAACATCGTGATGACATCGTTAAAGGCGCTGAAGGCCAGACTTAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAATCACCCCGAGCAAAAAACATCGTGATGACATCGTTAAAGGCGCTGAAGGCCAGACTTAC 540

Qy   541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCCGGACGCGCGGTGGCGGACGCGCGGTTCTTCA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCCATCTTTGTAAAGGCAGTTGCCGGTGGTGGCCGGACGCGCGGTGGCGGACGCGCGGTTCTTCA 600

Qy   601 CCTGATGAGCTTCCGCAAATTGGCAACAGAGACATCTCGTGAAGCGGCATTCGGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCTGATGAGCTTCCGCAAATTGGCAACAGAGACATCTCGTGAAGCGGCATTCGGC 660

Qy   661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720

Qy   721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGCTCACTCAGCGT 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 2401 ATTGGTTGCTGCATTCGCGCACACCCGGTTGGCGGCTGTTTCT 2460
         |||||||||||||||||||||||||||||||||||||||||||

Qy 2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGGACTGTGTACCTGCCATTTGAGTCTGGAACC 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGGACTGTGTACCTGCCATTTGAGTCTGGAACC 2520

Qy 2521 CCAGGCCCCAACCGGTCGCGTCTACCGCCACGACAGTGTCCAACCTG 2580
         ||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 CCAGGCCCCAACCGGTCGCGTCTACCGCCACGACAGTGTCCAACCTG 2580

Qy 2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 CGTGCACAGGCCACCGCACTGGGCCTTGCTGCTGATCGCTTCGAGCTCATCGAAGACAACTAC 2640

Qy 2641 GCAGCCGTTAATGAGATGGATGCGCGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 GCAGCCGTTAATGAGATGGATGCGCGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700

Qy 2701 GGCGACTTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCGACAC 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 GGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGCGACAC 2760

Qy 2761 CCACAAAGTACGACATCCCAGACTCGTCATCGCGTTCCTGCGCGAGCTTGGTAAC 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 CCACAAAGTACGACATCCCAGACTCGTCATCGCGTTCCTGCGCGAGCTTGGTAAC 2820

Qy 2821 CCTCCAGGTTGCTGGCCAGAACCACTCGCCACCCGCACACCGCGAGGGCGCCTCCGGAAGGC 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 CCTCCAGGTTGCTGGCCAGAACCACTCGCCACCCGCACACCGCGAGGGCGCCTCCGGAAGGC 2880

Qy 2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCCACTCGACGCTGATGATTCC 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAGAGCAGGCGCCACTCGACGCTGATGATTCC 2940

Qy 2941 AAGGAACGTGCAACAGCTGCTCCAACCGCCTCGCTGCGTCCCGAAGCCAACCGAAGAGTTCCTC 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 AAGGAACGTGCAACAGCTGCTCCAACCGCCTCGCTGCGTCCCGAAGCCAACCGAAGAGTTCCTC 3000

Qy 3001 GAGCACCGTCGCCGCTTCGGCAACCCTCGGCTGGATGATCGTTGCGCTGGATGATCGTTCGGACGGA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 GAGGCACCGTCGCCGCTTCGGCAACCCTCGGCTGGATGATCGTTGCGCTGGATGATCGTTCGGACGGA 3060

Qy 3061 CTGGTCGCGGAGGGCCGCCGCGAGACTTTGATCCGCGAGATGTGCCAGATGTGCCACCCCACTGCTTGTT 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 CTGGTCGCGGAGGGCCGCCGCGAGACTTTGATCCGCGAGATGTGCCAGATGTGCCACCCCACTGCTTGTT 3120

Qy 3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGTGGCCAACGTC 3180

Qy 3181 AACCGGCCAGATCCGCCCAATGCGTGTGCGGTGACCGCTCGTTGAGTCTGTCACCGCAACC 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 AACCGGCCAGATCCGCCCAATGCGTGTGCGGTGACCGCTCGTTGAGTCTGTCACCGCAACC 3240

Qy 3241 GCAGAAAAGGCAGATTCCTCCAACAAGGCCATGTTGCTGCAATGTTGGTGCCACATGTTGTC 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GCAGAAAAGGCAGATTCCTCCAACAAGGCCATGTTGCTGCAATGTTGGTGCCACATGTTGTC 3300

Qy 3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAGTCGCAATCATCGAG 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGAGATGCAGTCGCAATCATCGAG 3360

Qy 3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGCGTTGTG 3420

Qy 3421 GTTCCTGCTCAACGAAGGTGCAACGAAGGTGGAACGTGGCGACTTGATCGTCGTTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GTTCCTGCTCAACGAAGGTGCAACGAAGGTGGAACGTGGCGACTTGATCGTCGTTTCCTAA 3474

RESULT 2

```
Db   721  CTTCGCGATCGCACTGAGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT  780

Qy   781  CGTCACCAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCAGAACTGCGTGAT    840
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   781  CGTCACCAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCAGAACTGCGTGAT    840

Qy   841  CGCATTTGTTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CGCATTTGTTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC   900

Qy   901  GTGGAATTCTTGGTCGATGAAAAGGCAACCACGTTTCATCGAAATGAACCCACGTATC    960
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   901  GTGGAATTCTTGGTCGATGAAAAGGCAACCACGTTTCATCGAAATGAACCCACGTATC    960

Qy   961  CAGGTTGAGCACACCCGTGACTGAAGAGTCACCGAGGTGGACCTGGTCGAAGGCGCAGATG   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CAGGTTGAGCACACCCGTGACTGAAGAGTCACCGAGGTGGACCTGGTCGAAGGCGCAGATG   1020

Qy  1021  CGCTTGGCTGCTGGTGCAAACTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC   1080
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  1021  CGCTTGGCTGCTGGTGCAAACTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC   1080

Qy  1081  CACGGTGCAGCACTGCAGTGCCCGATCACCACCGAAGATCCAAACACACGGCTTCCGCCCA   1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  CACGGTGCAGCACTGCAGTGCCCGATCACCACCGAAGATCCAAACACACGGCTTCCGCCCA   1140

Qy  1141  GATACCGGAACTATCACCGCGTACCACCGGTACGCGTCCGCGTTCGTCTTGACGGT   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  GATACCGGAACTATCACCGCGTACCACCGGTACGCGTCCGCGTTCGTCTTGACGGT   1200

Qy  1201  GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTCGTGAAAATGACC   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTCGTGAAAATGACC   1260

Qy  1261  TGCCCGTGGTTCCGACTTTGAAACTGCTGTCGTGCACAGCGCCGGTTGGCTGAGTTC   1320
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db  1261  TGCCCGTGGTTCCGACTTTGAAACTGCTGTCGTGCACAGCGCCGGTTGGCTGAGTTC   1320

Qy  1321  ACCGTGTCTGTGTTGCAACCAACATTGGTTTCTTGCGTCGCGTTGCTGCGGGAAGAGAC   1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  ACCGTGTCTGTGTTGCAACCAACATTGGTTTCTTGCGTCGCGTTGCTGCGGGAAGAGAC   1380

Qy  1381  TTCACTTCCAAGCCGCATCGCCACCGGATTTATCGGCACCACACCTCCTTCAGGCT   1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TTCACTTCCAAGCCGCATCGCCACCGGATTTATCGGCACCACACCTCCTTCAGGCT   1440

Qy  1441  CCACCTGCCGGATGAGATGAGCAGGAGGACGCATCCTGGACAGGATGTCACCGTGAAC   1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  CCACCTGCCGGATGAGATGAGCAGGAGGACGCATCCTGGACAGGATGTCACCGTGAAC   1500

Qy  1501  AAGCCTCATGGTGCGTCGTCCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  AAGCCTCATGGTGCGTCGTCCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC   1560

Qy  1561  AAGGAATCTGCCACTGCCACGGCGGTTCCCGTGACCGGCTGGAAGCTTGGCCCCAGCCGCGC   1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  AAGGAATCTGCCACTGCCACGGCGGTTCCCGTGACCGGCTGGAAGCTTGGCCCCAGCCGCGCG   1620

Qy  1621  TTTGCTCGTGATCTCCGTGAGCTCCGGTGAGCAGGACGCACTGGCAGTTACTGCATCACCACCTTCCGCGAT   1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  TTTGCTCGTGATCTCCGTGAGCTCCGGTGAGCAGGACGCACTGGCAGTTACTGCATCACCACCTTCCGCGAT   1680

Qy  1681  GCACACCAGTCTTTGCTTGCTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGAACAGAG   1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  GCACACCAGTCTTTGCTTGCTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGGCGCAGAG   1740

Qy  1741  GCCGTGCAAAGTCGACTCCTGAGCTTTTGTCCGTGGAGGCGCCTGGGCGGGCGCGACCTAC   1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  GCCGTCGCAAAGTCGACTCCTGAGCTTTTGTCCGTGGAGGCGCCTGGGCGGGCGCGACCTAC   1800

Qy  1801  GATGTGGCGATGGCGTTTCCTCTTTTGAGGATTCCGTGGGACCGGGCTCGACGACTGGCGCGAG   1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  GATGTGGCGATGGCGTTTCCTCTTTTGAGGATTCCGTGGGACAGGCTCGACGAGCTGCGCGAG   1860
```

```
Qy  1861  GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCGCGGATCGGGATCGTGGGATACACC   1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCGCGGATCGGGATCGTGGGATACACC   1920

Qy  1921  CCGTACCCAGACTCCGTCTGCGCGCGGTTTGTTAAGGAAGCTGCCAGCTGCCAGCTCCGGCGTTGGAC   1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CCGTACCCAGACTCCGTCTGCGCGCGGTTTGTTAAGGAAGCTGCCAGCTGCCAGCTCCGGCGTTGGAC   1980

Qy  1981  ATCTTCCGCCATCTTGCACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA   2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ATCTTCCGCCATCTTGCACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA   2040

Qy  2041  GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT   2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT   2100

Qy  2101  GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC   2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  GATCCAAATGAAAAGTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC   2160

Qy  2161  AAGTCTGCGCGTCACATTCTGCCATTAAGGATATGCGTCGGTCTCGCTCGCCCAGCTGCG   2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  AAGTCTGCGCGTCACATTCTGCCATTAAGGATATGCGTCGGTCTCGCTCGCCCAGCTGCG   2220

Qy  2221  GTAACCAAGCTCGTCACCGACTGCGCCGGTGAATTCCATCTGCCAGTGCACTGCACACC   2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  GTAACCAAGCTGGTCACCGACTGCGCCGGTGAATTCCATCTGCCAGTGCACTGCACACC   2280

Qy  2281  CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTCGCGAGCTCAAGCTGGTGCAGAT   2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CACGACACTGCGGGTGGCCAGTTGGCCAGTTGGCTACCTACTTTGCTCGCGAGCTCAAGCTGGTGCAGAT   2340

Qy  2341  GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGCACCACTCCCTGTCTGCC   2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCTGTTGACGGTGCTTCCGCACCACTGTCTGCGCACCACTCCCTGTCTGCC   2400

Qy  2401  ATTGTTGCTGCAATTCGCGCAACCCGTGCGCGATACCGGTTTGAGCGTGTTTCT   2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  ATTGTTGCTGCAATTCGCGCAACCCGTGCGCGATACCGGTTTGAGCGTGTTTCT   2460

Qy  2461  GACCTCGAGCCGTACTGGGAAGCTGTGCGCCGACTGTACCTGCCATTTGAGTCTGGAACC   2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GACCTCGAGCCGTACTGGGAAGCTGTGCGCCGACTGTACCTGCCATTTGAGTCTGGAACC   2520

Qy  2521  CCAGGCCCAACCGTCGCGTCGCTCCACCACGCGAAATCCCAGGCGGACAGTTGTCCAACCTG   2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  CGTGCACAGGCCCAACCGTCGCGTCGCTCCACCACGCGAAATCCCAGGCGGACAGTTGTCCAACCTG   2580

Qy  2581  CGTGCACAGGCCACCCGACTCGGGCCTTGCTCGTCATCGATCGTGCTTCGAAGACAACTAC   2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CGTGCACAGGCCACCCGACTCGGGCCTTGCTCGTCATCGATCGTGCTTCGAAGACAACTAC   2640

Qy  2641  GCAGCCGGTTAATGACATGAGATGCGGATCTGGGACGCGCCAACCAAGTCACCCCATCCTCCAAGGTTGTT   2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  GCAGCCGGTTAATGACATGAGATGCGGATCTGGGACGCGCCAACCAAGTCACCCCATCCTCCAAGGTTGTT   2700

Qy  2701  GGCGACCTGCACTCCACCTGGTTGCGGGTGAGATCCAGCAGACTTTGCTGCGAGAC   2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  GGCGACCTGCACTCCACCTGGTTGCGGGTGAGATCCAGCAGACTTTGCTGCGAGAC   2760

Qy  2761  CCACAAAAGTACGACATCCACGACTCTGTCATCGCGTTCCTGCGCGGCACTTGGTAAC   2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  CCACAAAAGTACGACATCCACGACTCTGTCATCGCGTTCCTGCGCGGCACTTGGTAAC   2820

Qy  2821  CCTCCAGTGGCTGGCACGAACCACTGCCACACCGCCGCGAGGC   2880
          ||||||||||||||||||||||||||||||||||||||||||
Db  2821  CCTCCAGTGGCTGGCACGAACCACTGCCACACCGCCGCGAGGC   2880

Qy  2881  AAGGCACCTCTGACGAAGTTCCTGACGACGCACCGCACCTCGCGCGAG   2940
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  AAGGCACCTCTGACGAAGTTCCTGACGAAGCAGCGCACCTCGCGCGAG   2940
```

```
Qy  2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTCGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 AAGGAACGTCGCAACAGCCTCAACCGCCTGCTCGTTCCCGAAGCCAACCGAAGAGTTCCTC 3000

Qy  3001 GAGCACCGTGCCGCTTCGCCGCAACACCTCGCGCTGGATCGTGAATTCTTCTACGGA 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 GAGCACCGTGCCGCTTCGCCGCAACACCTCGCGCTGGATCGTGAATTCTTCTACGGA 3060

Qy  3061 CTTGGTCGAGGGCGCCGAGACTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 CTTGGTCGAGGGCGCCGAGACTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTT 3120

Qy  3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGGCCAACGTC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 CGCCTGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCAATGTTGGCCAACGTC 3180

Qy  3181 AACGGCCAGATCCGCCCCAATGCGTGCGTGCCGACCGCTCCGTTGAGTGCTGTCTGTT 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 AACGGCCAGATCCGCCCCAATGCGTGCGTGCCGACCGCTCCGTTGAGTGCTCGTCACCGCAACC 3240

Qy  3241 GCAGAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTC 3300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GCAGAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGCTGTTGTC 3300

Qy  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCCAATCATCGAG 3360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATGCAGTCGCCAATCATCGAG 3360

Qy  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCGGCAAGATTGAACGCGTTGTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGCGGCAAGATTGAACGCGTTGTG 3420

Qy  3421 GTTCCTGCTGCAACGAAGGTGGAAGTGGCGACTTGATCCTCGTCGTTCCTAA 3474
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GTTCCTGCTGCAACGAAGGTGGAAGTGGCGACTTGATCCTCGTCGTTCCTAA 3474
```

RESULT 3
US-10-045-072-1
; Sequence 1, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: SinsKey, Anthony J.
; APPLICANT: Lessard, Phillip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533 0790002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(3621)
US-10-045-072-1

```
Query Match           97.8%;  Score 3398.8;  DB 9;  Length 3621;
Best Local Similarity 98.6%;  Pred. No. 0;
Matches 3427;  Conservative   0;  Mismatches 47;  Indels   0;  Gaps   0;

Qy    1 GTGACTGCTATCACCCTTGCGCGGTCTCTTGTTGAAGGAATAATTACTCTAGTGTCGACT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  148 GTGACTGCTATCACCCTTGCGCGGTCTCTTGTTGAAGGAATAATTACTCTAGTGTCGACT 207

Qy   61 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGGAAATC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  208 CACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGGAAATC 267

Qy  121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAACTCGAAACCGGTGCAGCCACGGTAGCTATTACCCC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAACTCGAAACCGGTGCAGCCACGGTAGCTATTACCCC 327

Qy  181 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGATTGGTACT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  328 CGTGAAGATCGGGGATCATTCCACCGCTCTTTGCTTCTGAAGCTGTCCGATTGGTACC 387

Qy  241 GAAGGCTCACCAGTCAAGGCGGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  388 GAAGGCTCACCAGTCAAGGCGGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447

Qy  301 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 360
        ||||||||||||||| ||||||||||||||||||||||||||| ||||||||||||||
Db  448 GTTAAAGCAGATGCCATTTACCCGGATAACGGCTTCCTGTCTGTAAAATGCCCAGCTTGCC 507

Qy  361 CGCGAGTGCGCGGAAAACGCATTACTTTATTGGCCACCAGAGGTTCTTGATCTC 420
        |||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  508 CGCGAGTGTGCGGAAAACGCATTACTTTATTGGCCACCAGAGGTTCTTGATCTC 567

Qy  421 ACCGGTGATAAGTCTGTGCGTAACCGCCGGAGAGAGGCGGTCTGCCAGTTTGGCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  568 ACCGGTGATAAGTCTGTGCGTAACCGCCGGAGAGAGGCGGTCTGCCAGTTTGGCG 627

Qy  481 GAATCCACCCCGAGCAAAAACATCGATGACTCGTTAAAGGCGCTGAAGGCCAGACTTAC 540
        |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  628 GAATCCACCCCGAGCAAAAACATCGATGAGTCGTTAAAGGCGCTGAAGGCCAGACTTAC 687

Qy  541 CCCATCTTGTAAAGGCAGTTGCGCGGTGGTGGCGGCGATCGTGAAGCGCATCTCGTGAAGCGGCATTCGGC 600
        |||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  688 CCCATCTTTGTGAAGGCAGTTGCGCGGTGGTGGCGGCGATCGTGAAGCGCATCTCGTGAAGCGGCTTTCGGC 747

Qy  601 CCTGATGAGCTCCGCAAATTGCAACAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 660
        |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db  748 CCTGATGAGCTTCGCAAATTAGCAACAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGGC 807

Qy  661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATC 720
        ||| ||  |||||||||||||||||||||||||||||||| ||| ||||||||||||||
Db  808 GATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867

Qy  721 CTTGGCGATCGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
        |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927

Qy  781 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCGTGAT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  928 CGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTGGATCCAGAACTGCGTGAT 987

Qy  841 CGCATTGTGCGGATGCAGTAAAGTTCTCGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  988 CGCATTGTGCGGATGCAGTAAAGTTCTCGCCGCTCCATTGGTTACCAGGGCGCGGGAACC 1047

Qy  901 GTGGAATTCTTGGTCGATGATGAAAAGGGCAACCACGTTCATCGAAATGAACGAAATCGTATC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db 1048 GTGGAATTCTTGGTCGATGATGAAAAGGGCAACCACGTTCATCGAAATGAACCACGTATC 1107

Qy  961 CAGGTTGAGCACACCGTGACTGGAAGGCGCAGATCCGAGTGGACTGGAGCGCAGATG 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db 1108 CAGGTTGAGCACACCGTGACTGGAAGGCGCAGATCCGAGTGGACTGGAGGCGCAGATG 1167

Qy 1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGCACCCGAGCTAAGATAAGCAGACC 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGCACCCGAGCTAAGATAAGCAGACC 1227

Qy 1081 CAGGTGCAGCACGACTGCAGTCATCACCACGAGATCCAAACAACGGCTTCCGCCCA 1140
        || ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1228 CACGGTGCAGCACGACTGCAGTCATCACCACGAAGATCCAAACAACGGCTTCCGCCCA 1287

Qy 1141 GATACCGGAACTATCACCGCTACCCGTCACCAGGGCGAGCTGGCGTTCGTCTTGACGGT 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2281 CACGACACTGCGGGTGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2340
          ||||||||||||||  |||||||||||| |||||||||||||||||||||||||||||||
Db   2428 CACGACACTGCGGGTGCCAGCTGGCAACCTCACTTTGCTGCAGCTCAAGCTGGTGCAGAT 2487

Qy   2341 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCACCACTGTCCGCACCACTCCCTGCTCGCC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2488 GCTGTTGACGGTGCTTCCGCACCACTGTCTGCACCACTGTCCGCACCATCCCTGCTCGCC 2547

Qy   2401 ATTGTTGCTGCATTCGCGCACCCGTCGCGCACACCGGTTGAGCCTCGAGGCTGTTTCT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2548 ATTGTTGCTGCATTCGCGCACCCGTCGCGGATACCGGTTTGAGCCTCGAGGCTGTTTCT 2607

Qy   2461 GACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACC 2520
          ||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db   2608 GACCTCGAGCCGTACTGGGAAGCAGTGCGCGGGACTGTACCTGCCATTTGAGTCTGGAACC 2667

Qy   2521 CCAGGGCCCAACCGGTGCGTGCTGCCCGTCTACCGCCACGGAATCCCAGGCGGACAGTTGTCCAACCTG 2580
          ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2668 CCAGGGCCCAACCGGTGCGTGCTGCCCGTCTACCGCCACGGAATCCCAGGCGGACAGTTGTCCAACCTG 2727

Qy   2581 CGTGCACAGGGCCACCGCACTGGGCCTTGCTGATCGGTCGCTCATCGAGACACAACTAC 2640
          ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db   2728 CGTGCACAGGGCCACCGCACTGGGCCTTGCTGATCGGTTGCGAACTCGTTCGAAGCAACAACTAC 2787

Qy   2641 GCAGCCGTTAATGAGATGCTGGGACGGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2788 GCAGCCGTTAATGAGATGCTGGGACGGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT 2847

Qy   2701 GGCGACTTCGCACTCCACCTGGTTGGTGCGGGTGAGTCAGCAGACTTTGCTGCAGAC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   2848 GGCGACTTCGCACTCCACCTGGTTGGTGCGGGTGTGTGGATCCAGCAGAGACTTTGCTGCCGAT 2907

Qy   2761 CCACAAAAGTACGACATCCAGACTCCGTCCATCGCGGTTCCTGCGGCGGCGAGCTTGGTAAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2908 CCACAAAAGTACGACATCCAGACTCCGTCCATCGCGGTTCCTGCGGCGGCGAGCTTGGTAAC 2967

Qy   2821 CCTCCAGGTGGCTGGCCAGAGAAGTTCCTGGAAGAGCAGGCGCACCTCGACGCGTGATGATTCC 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2968 CCTCCAGTGGCTGGCCAGAGACGCCACTGCGCGCACCGCGGCGGCACTGGACGCGTCCGAAGGC 3027

Qy   2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCTGCACCTCGAGCGCTGATGATTCC 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3028 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCTGCACCTCGAGCGCTGATGATTCC 3087

Qy   2941 AAGGAACGTCGCAACAGCCTCGTCCGCTGTTCCCGAGCCAACCGAGGAGTTCCTC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3088 AAGGAACGTCGCAATAGCCTCACCCTGCTGTTCCCGAGCCAACCGAGGAGTTCCTC 3147

Qy   3001 GAGCACCCGTCGCCGCTTCGGCAACACCTCTGCGCCTGCGATGATCGTGAATTCTTCTACGGA 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3148 GAGCACCCGTCGCCGCTTCGGCAACAACTCTGCGCCTGCGATGATCGTGAATTCTTCTACGGC 3207

Qy   3061 CTGGTGGAGGGCCGGCGCCTTAACGCTCTCCCAGCAGTGCGTCCGCCAGCAGTGTGCGCAACCGCTGTT 3120
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 CTGGTCGGAGGGCCGGCGCCTTAACGCTCTCCCAGGAGACTTTGATCCGCCAGCAGTGTGCGCAACCGTC 3267

Qy   3121 CGCTCGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCCAACGTGTGTGGCCAACGTC 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 CGCTCGGATGCGATCTCTGAGCCAGACGATAAGGGTATGCGCCAATGTGTGTGGCCAACGTC 3327

Qy   3181 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 AACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387

Qy   3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTACAAGGGTCAGGCTGGAATCGAG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTACAAGGGTCAGGCTGGAATCATCGAG 3447

Qy   3301 ACTGTGACTGTTGCTGAAGGTGATGAGGTCAATCGAG 3360
          || ||||||||||||||||||||||||||||||||||
Db   3448 ACCGTGACTGTTGCTGAAGGTGATGAGGTCAATCATCGAG 3507
```

```
Db   1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGCGGAGCTGGCGTTCGTCTTGACGGT 1347

Qy   1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCGTCGTGAAAATGACC 1260
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCGTCGTGAAAATGACC 1407

Qy   1261 TGCCCTGGTTCCGACTTTGAAACTGCTCGTTGCTGTGCACAGCGCGCGTTGGCTGAGTTC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 TGCCCTGGTTCCGACTTTGAAACTGCTCGTTGCTGTGCACAGCGCGGGTTGGCTGAGTTC 1467

Qy   1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCCGTTGCTGCGGAAGAGGAC 1380
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCCGTGGCTGCGGGGAGAGGAC 1527

Qy   1381 TTCACTTCCAAGGCGGCATCGCCACCGGATTATCGGCGATCACCCACACCTTCCTTCAGGCT 1440
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db   1528 TTCACTTCCAAGGCGGCATCGCCACCGGATTCATTGCCGATCACCCACACCTTCCTTCAGGCT 1587

Qy   1441 CCACCTGCGGATGATGAGCAGGCATCGATTACTTGGCAGATGTCACCGTGAAC 1500
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   1588 CCACCTGCTGATGATGAGCAGGCATCGATTACTTGGCAGATGTCACCGTGAAC 1647

Qy   1501 AAGCCTCATGGTGTCGTCCAAAGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC 1560
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||| |||
Db   1648 AAGCCTCATGGTGTCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707

Qy   1561 AAGGATCTGCCACTGCCACGCGGTTCCCGTGAAGCAGCTTGGCCTGCAGCCGCCG 1620
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1708 AAGGATCTGCCACTGCCACGCGGTTCCCGTGAAGCAGCTTGGCCTGCAGCCGCCG 1767

Qy   1621 TTTGCTCGTGATCTCCGTGAGCAGCAGCGCACTGCCAGTTACTGATACCACCTTCCGCGAT 1680
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1768 TTTGCTCGTGATCTCCGTGAGCAGCAGCGCACTGCCAGTTACTGATACCACCTTCCGCGAT 1827

Qy   1681 GCACACCAGTCTTTTGCTTCGACCCGAGTCGCTCATTCGCGCTGAAGCCTGCGGCGCAGAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 GCACACCAGTCTTTTGCTTCGACCCGAGTCGCTCATTCGCGCTGAAGCCTGCGGCGCAGAG 1887

Qy   1741 GCCGTCGCAAAGCTGACTCCTGACTTTTGTCCGTGGAGGCTTGGGAGGCCGGGCGGCGACCTAC 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1888 GCCGTCGCAAAGCTGACTCCTGACTTTTGTCCGTGGAGGCTTGGGAGGCCGGGCGGCGACCTAC 1947

Qy   1801 GATGTGGCGATGCGTTCCTCTTTTGAGGATCCGTTGGGACCAGGCTGCACGAGCTGCGCGGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1948 GATGTGGCGATGCGTTCCTCTTTTGAGGATCCGTTGGGACCAGGCGTCCACGAGCTGCGCGGAG 2007

Qy   1861 GCGATGCCGAAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTGGGATACACCC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTGGGATACACCC 2067

Qy   1921 CCGTACCCAGACTCCGTCTGCCCGTCGTTTGTTAAGGAAGCTGCCCAGCAATCGACGCA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2068 CCGTACCCAGACTCCGTCTGCCCGTCGTTTGTTAAGGAAGCTGCTCCCAGCAATCGACGCA 2127

Qy   1981 ATCTTCCGCATCTTCGACGCGCCTTAACGCGTGCTTCCCAGAGTGCGTCAGCAATCGACGCA 2040
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2128 ATCTTCCGCATCTTCGACGCGCCTTAACGCGTGCTTCCCAGAGTGCGTCAGCAATCGACGCA 2187

Qy   2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCGTGTGATCGTCT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCGTGTGATCGTCTCT 2247

Qy   2101 GATCCAAAATGAAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGAGATCGTC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2248 GATCCAAAATGAAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGAGATCGTC 2307

Qy   2161 AAGTTCTGGCGCTCACATTCTCGGCCATTAAGGATATGGCTGGTTCTGCTTCGCCCAGCTGCG 2220
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   2308 AAGTTCTGGCGCTCACATTCTCGGCCATTAAGGATATGGCTGGTTCTGCTTCGCCCAGCTGCG 2367

Qy   2221 GTAACCAAGCTGGTCACCCGACCAGTGCCAGTCGCACGTGCACACC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2368 GTAACCAAGCTGGTCACCCGACCAGTGCCAGTCGCACGTGCACACC 2427
```

```
Qy   3361  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGGCGTTGTG  3420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508  GCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAACGGCGTTGTG  3567

Qy   3421  GTTCCTGCTGCTCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTTCCTAA  3474
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568  GTTCCTGCTGCTCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTTCCTAA  3621
```

RESULT 4
US-09-738-626-765
; Sequence 765, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 765
; LENGTH: 3420
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-765

Query Match   96.3%; Score 3344.8; DB 9; Length 3420;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy    52  GTGTCGACTCACACATCTTCAAGCGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC   111
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||| ||||||
Db     1  GTGTCGACTCACACATCTTCAAGCGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGC    60

Qy   112  GGGCAAATGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCT    171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GGGCAAATCGCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGTAGCT   120

Qy   172  ATTTACCCCGTGAAGATCGGGGATCATTCCACGGTCTCTTTGCTTCTGAAGCTGTCCCGC   231
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  ATTTACCCCGTGAAGATCGGGGATCATTCCACGGTCTCTTTGCTTCTGAAGCTGTCCCGC   180

Qy   232  ATTGGTACTGAAGGCTCACCAGTCAAGGCGTCAAGGCGTTACCTGGACATCGATGAAATTATCGGTGCA   291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ATTGGTACCGAAGGCTCACCAGTCAAGGCGTCAAGGCGTTACCTGGACATCGATGAAATTATCGGTGCA   240

Qy   292  GCTAAAAAGTTAAAGCAGATGCTATTTACCCGGGATATATGGCTTCCTGTCTGAAAATGCC   351
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   241  GCTAAAAAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCC   300

Qy   352  CAGCTTGCCCGCGAGTAGCTCGTGCGGTAACCGCCGGAAGGCTGGTCTGCCA   411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CAGCTTGCCCGCGAGTGCGGTAACCGCCGGAAACCCCCAGAGGTT   360

Qy   412  CTTGATCTCACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAAGGCTGGTCTGCCA   471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  CTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGGAAGGCTGGTCTGCCA   420
```

```
Qy   472  GTTTTGGCGGAATCACCCCGAGCAAAAAACATCGATGACATCGTTAAAAGCGCTGAAGGC   531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GTTTTGGCGGAATCACCCCGAGCAAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGC   480

Qy   532  CAGACTTACCCCATCTTGTAAAGGCAGTTGCCGGTGGTGGGACGCGGTATGCGCTTT   591
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CAGACTTACCCCATCTTGTGAAGGCAGTTGCCGGTGGTGGGACGCGGTATGCGCTTTT   540

Qy   592  GTTTCTTCACCTGATGAGCTTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCG   651
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCG   600

Qy   652  GCATTCGGCGACGGTTCGGTATATGTCGAACGTGCTGATTAACCCCAGCACATTGAA   711
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   601  GCTTTCGGCGATGGCGGGTATATGTCGAACGTGTGATTAACCCTCAGCATATTGAA   660

Qy   712  GTGCAGATCCTTGGGCATCGACTGAGAAGTGTACACCTTATGAACGTGACTGCTCA   771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GTGCAGATCCTTGGGCATCACACTGAGAGAGTGTACACCTTATGAACGTGACTGCTCA   720

Qy   772  CTTGACGCGTCGTCACCAAAAAGTTGTCGAAATTGCGGCCAGCACAGCAGCATTTGGATCCAGAA   831
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CTTGACGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCAGCATTTGGATCCAGAA   780

Qy   832  CTGCGTGATCGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCGCCTCCATTGGTTACCAGGGGC   891
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CTGCGTGATCGCATTTGTGCGGATGCAGTCAGTAAAGTTCTGCGCCTCCATTGGTTACCAGGGGC   840

Qy   892  GCGGGAACCGTGGAATTCTTGTCGATGAAAAGGGCAACCAGTTTTCATCGAAATGAAC   951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAAC   900

Qy   952  CCACGTATCCAGGTTGAGCACACCGTGACTGAGAAGAGTCACCGAGGTGGACCTGGTGAAG   1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  CCACGTATCCAGGTTGAGCACACCGTGACTGAGAAGAGTCACCGAGGTGGACCTGGTGAAG   960

Qy  1012  GCGCAGATGCGCTTGGCTGCTGCGCTGCAGATCGTCGAGTGCGCATCACCACGGAAGATCCAAACAACGGC   1071
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GCGCAGATGCGCTTGGCTGCTGCGCTGCAGATCGTCGAGTGCGCATCACCACGGAAGATCCAAACAACGGC   1020

Qy  1072  ATCAAGACACCACGGTGCAGTCGAGTGCCGCATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT   1131
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ATCAAGACACCACGGTGCAGTCGAGTGCCGCATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGT   1080

Qy  1132  TTCCGCCAGATACCGGAACTATCACCGGTGCGGAATCCGGTATCACCGGTATCACCGGCGGAGCTGGCGTTCGT   1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TTCCGCCAGATACCGGAACTATCACCGGTGCGGAATCCGGTATCACCGGTATCACCGGCGGAGCTGGCGTTCGT   1140

Qy  1192  CTTGACGGTGCAGCTCAGCTCCGGTGCGAATCCGGTGCAGCCACGTAGCT   1251
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CTTGACGGTGCAGCTCAGCTCCGGTGCGAATCCGGTGCAGCCACGTAGCT   1200

Qy  1252  AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGTCGTGTGCTGCTGTGGTTCTCCGAAGCTGTCCCG   1311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  AAAATGACCTGCCGTGGTTCCGACTTTGAAACTGTCGTGTGCTGCTGTGGTTCTCCGAAGCTGTCCGC   1260

Qy  1312  GCTGAGTTCACCGTGTCGTGTTGCAACCAACATTGGTTCTTCGGTGTGCTGCGCG   1371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GCTGAGTTCACCGTGTCGTGTTGCAACCAACATTGGTTCTTCGGTGTGCTGCGG   1320

Qy  1372  GAAGAGAGACTTCACTTCCAAGCGCATCGCCACCGGAATCGCCGGGATTATCGGCGATCACCACACCTC   1431
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  GAAGAGAGACTTCACTTCCAAGCGCATCGCCACCGGAATTATTGCCGGATCACCCGCACCTC   1380

Qy  1432  CTTCAGGCTCCACCTGCGGATGAGATGAGCAGGGACCGGCATCCTGGCACCGAGAGTC   1491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  CTTCAGGCTCCACCTGCGGATGAGATGAGCAGGGACCGGCATCCTGGCAGATGTC   1440

Qy  1492  ACCGTGAACAAGCTTCATGTGTCGGTCCAAAGAATGTTGCAGCACCAATCGATAAGCTG   1551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  ACCGTGAACAAGCTTCATGTGTCGGTCCAAAGAATGTTGCAGCTCCTATCGATAAGCTG   1500
```

```
Qy   1552 CCCAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAGAAGCAGCTTGGC 1611
          || |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db   1501 CCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAGCAGCAGCTTGGC 1560

Qy   1612 CCAGCCGGCGGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAGCCGGCGGTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACC 1620

Qy   1672 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1731
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCT 1680

Qy   1732 GCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGGCGGC 1791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCGGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGGCGGC 1740

Qy   1792 GCGACCTACGATGTGGCGATGCCGTTCCTCTTTGAGGATCCGTGGGACCAGGCTCGACGAG 1851
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GCGACCTACGATGTGGCGATGCCGTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAG 1800

Qy   1852 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTG 1911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCGCAACACCGTG 1860

Qy   1912 GGATACACCCCGTACCCAGACTCCGTCTGCGCGGTTTGTTAAGGAAGCTGCCAGCTCC 1971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GGATACACCCCGTACCCAGACTCCGTCTGCGCGGTTTGTTAAGGAAGCTGCCAGCTCC 1920

Qy   1972 GGGGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGGCTCCAGCA 2031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGCGTGGACATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGGCTCCAGCA 1980

Qy   2032 ATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATTCTTGGT 2091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATCGACGCGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATTCTTGT 2040

Qy   2092 GATCTCCTGATCCAAATGAAAAGCTCTACACCCGGATTACTACCTAAAGATGGCAGAG 2151
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GATCTCCTGATCCAAATGAAAAGCTCTACACCCGGATTACTACCTAAAGATGGCAGAG 2100

Qy   2152 GAGATCGTCAAGTCTGGCCTCACATTCTGGCCATTAAGGATATGGCTCGTCTCTTCGC 2211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GAGATCGTCAAGTCTGGCCTCACATCTTGGCCATTAAGGATATGGCTCGTCTCTTCGC 2160

Qy   2212 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCCCGTGAATTCGATCTGCCAGTGCAC 2271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCCCGTGAATTCGATCTGCCAGTGCAC 2220

Qy   2272 GTGCACACCCCAGACGACTGCGCGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCT 2331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GTGCACACCCCAGGACGACTGCGCGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCAAGCT 2280

Qy   2332 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCC 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTCGGCACCACCTCCCAGCCATCC 2340

Qy   2392 CTGTCTGCCATTGTTGCTGCATTCGCGCACCCGTCGCGGATACCGGTTTGAGCCTCGAG 2451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CTGTTCTGCCATTGTTGCTGCATTCGCGCACCCGTCGCGGATACCGGTTTGAGCCTCGAG 2400

Qy   2452 GCTGTTTCTGACCTCGAGCCGTACCGGAAGCTGTGCGCGCGACTGTACCTGCCATTGAG 2511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GCTGTTTCTGACCTCGAGCCGTACCGGAAGCAGCGGCGGACTGTACCTGCCATTGAG 2460

Qy   2512 TCTGGAACCCCAGGCCCAACCGCTTCCACCGCTCGCGGGCTTCACCGGGACAGTTG 2571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TCTGAACCCCAGGCCCAACCGCCAACCGCGTCTACCGCCACCGCAGGGGACAGTTG 2520

Qy   2572 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCTGATCGCTTCGAGCTCATCGAA 2631
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGATGTTTCGAACTCATCGAA 2580

Qy   2632 GACAACTACGCAGCCGGTTAATGAGATGCTGGACGCCCAACCAAGGTCACCCCATCCTCC 2691
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GACAACTACGCAGCCGGTTAATGAGATGCTGGACGCCCAACCAAGGTCACCCCATCCTCC 2640

Qy   2692 AAGGTTGTTGGGCACCTCGCACTCCACCTCGGGGTCGGGGTGTAGATCCAGCAGACTTT 2751
          ||||||||||||||||||||||||| |||||||  || |||||||||||||||||||||
Db   2641 AAGGTTGTTGGGCACCTCCGCACTCCACCTCGGTGGTGGGGTGTGATCCAGCAGACTTT 2700

Qy   2752 GCTGCAGACCCACAAAAGTACGACATCCCAGACTCGTCATCGCGTTCCTGCGCGGCGGAG 2811
          ||||| ||| |||||||| ||||||||||||||||||| |||||||  ||||||||||||
Db   2701 GCTGCCGGATCCACAAAAGTACGACATCCCAGACTCGTCATCGCGTTCCTGCGCGGCGGAG 2760

Qy   2812 CTTGGTAACCCTCCAGGTGGCTGGCCAGAACCACTGCGCACCTGCGCACCTGGAAGGCCGC 2871
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   2761 CTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCTGCGCACCTGGAAGGCCGC 2820

Qy   2872 TCCGAAGGCAAGGCACCTCTGACGGAGTTCCTGACGGAGAAGAGCAGGCGCACCTCGACGCT 2931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TCCGAAGGCAAGGCACCTCTGACGGAGTTCCTGACGGAGAGAGCAGGCGCACCTCGACGCT 2880

Qy   2932 GATGATTCCAAGAACGTCGCAACAGCCTCAACGCCTGTGTTCCCGAAGCCAACCGAA 2991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GATGATTCCAAGAACGTCGCAATAGCCTCAACCGCTGTGTTCCCGAAGCCAACCGAA 2940

Qy   2992 GAGTTCCTGAGCACCGTCGCCCTTCGCCAACACCTTCGCGCTGATGATGATCGTGAATTC 3051
          ||||||| ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   2941 GAGTTCCTCGAGCACCGTCGCCCTTCGCCAACACCTTCCGCCGCTGATGATGATCGTGAATTC 3000

Qy   3052 TTCTACGGACTGTCGAGGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCA 3111
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   3001 TTCTACGGACTGTCGGTCGAGGCCGCGAGACTTTGATCCGCCCTGCCAGATGTGCGCACCCCA 3060

Qy   3112 CTGCTTGTTCGCCTGGATCGAGCTCTGAGCCGACAGATAAGGGTATGCGCAATGTTGTG 3171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CTGCTTGTTCGCCTGGATCGAGCTCTGAGCCGACAGATAAGGGTATGCGCAATGTTGTG 3120

Qy   3172 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGCTCCGCTCCGTTGAGTCTGTC 3231
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGCTCCGCTCCGTTGAGTCTGTC 3180

Qy   3232 ACCGCAACCGCAGAAAAGCAGATTCCTCCAACAAGGCCATGTTGCTGCACCATTCGCT 3291
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 ACCGCAACCGCAGAAAAGCAGATTCCTCCAACAAGGCCATGTTGCTGCACCATTCGCT 3240

Qy   3292 GGTGTTGTCACTTGTGACTGTTGCTGAGGTGATGAGCTCAAGGCTGCGAGGATGCAGTCGCCA 3351
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   3241 GGTGTTGTCACCTGTGACTGTTGCCGTGAGGTGATGAGGCTCAAGGCTGCGAGGATGCAGTCGCA 3300

Qy   3352 ATCATCGAAGGCTATGAAGATGAAGCAACAATCACTGCTTCTGTTGACGGCAAGATTGAA 3411
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 ATCATCGAGGCTATGAAGATGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGAT 3360

Qy   3412 CGGCGTTGGTTCCTGCTCGCAACGAAGGTGGAAGGTGCGCACGACTTGATCGTCGTTTCC 3471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 CGGCGTTGGTTCCTGCTCGCAACGAAGGTGGAAGGTGCGCACGACTTGATCGTCGTTTCC 3420
```

RESULT 5
US-09-917-800A-1566
; Sequence 1566, Application US/09917800A
; Patent No. US2002/0119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31

```
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR APPLICATION FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR APPLICATION FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR APPLICATION FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR APPLICATION FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR APPLICATION FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR APPLICATION FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR APPLICATION FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR APPLICATION FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1566
; LENGTH: 3945
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012744
US-09-917-800A-1566

Query Match     19.6%;   Score 681;  DB 10;  Length 3945;
Best Local Similarity   54.2%;   Pred. No. 1.3e-200;
Matches 1521; Conservative   0; Mismatches 1260;  Indels  24;  Gaps  6;
```

```
Qy   86  TCAAAAAGATCTTGTAGCAAACCGGCGGGCGAAATCGCGGTCCGTTGCTTTCCGTGCAGCAC  145
         |||||                ||||       ||| ||     |   |||
Db  144  TCAAGAAAGTGTGGCCAACAGAGTGAGATTGCCATCCGAGTGTTTCGTGCTGCTGCA     203

Qy  146  TCGAAACCGGTGCGAGCACCGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCACC    205
         |||||         ||||
Db  204  CAGAGCTGGGTATCCGCACAGTGCCTGCTCTACTCGGAGACAGGACACAGGACACC      263

Qy  206  GCTCTTTGCTCTGAAGCTGTCCCATTGGTACTGAAGGCTCACCGTCAAGGCGTACC      265
         ||     ||||||||             |||
Db  264  GGCAGAAAGCTGATGAAGCTACCTTATTCGGCCGTGGGCGTCCTCGTGTCCAAGCCTACC  323

Qy  266  TGGACATCGATGAAATTATCGGTCGCAGCTAAAAAAGTTAAAGCAGAGATGCTATTTACCCCG  325
         ||    || ||||||
Db  324  TGCACATTCCAGACACATTAAGGTGCCAAGGAGAGTGTAGATGCTGTGCACCCTG       383

Qy  326  GATATGGCTTCCTGTCTGAAAATGCCAGCTTGCCCGAGTGCGCGGAAAACGGCATTA      385
         ||||
Db  384  GCTATGGGTTCCTCGAGAGAGCAGAGGTTCTCCCACGGCCTGCCCAAGATGCTGGAGTCC    443

Qy  386  CTTTTATTGCCCAACCCAGAGTTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTAA      445
         ||||||                 |||
Db  444  GATTCATTGGTCCCAGAGCCCAGAGGTGGTCCGCAAGATGGGAGACAACAAGGTGGAAGCCGGGG  503

Qy  446  CCGCCGCGGAAGAAGCGTCGTGCTCGCCAGTTTGGGCGGAATCCA---CCCCGAGCAAAAAACA  502
         ||||                                            ||| ||||||||
Db  504  CCAATGCCATTGCCATTGCTTCAGCGGCGTTCCCATGTCCCACTAATTCCCCACTCAATTCCC    563

Qy  503  TCGATGACATCGTTAAAAGCGCTGAAGCGACTTACCCCATCTTTGTAAAGGCAGTTG        562
         |||                   ||||
Db  564  CCCCAGCTGGTTTCAGAGAGACATCCTTCTGCGGACATCCTTCTCCTATATTCTTCAAGGCTGCCT  623

Qy  563  CCGGTGGTGGCGGACGCGGTATGCCGTTTGTTCTTTCTTCACCTGATGAGCTCCGCAAATTGG   622
         ||                                           |||
Db  624  ATGGGACCACCAGGGGGCCGTGCATGAGCGGCTTGTGTCATGAGGAGCTGGAAGAGAATT      683

Qy  623  CAACAGAAGCAGTCTCGTGAAGCTGAAGCGGCATTCGGCCGACGGTTCGGTATATTGTCGAAC   682
         ||                                  |||
Db  684  ACACCCCGACCTCACCCTGAGGCCTTGCGGGCCTTTGGGAATGGGCATTGTTTTGTGGGAGA    743

Qy  683  GTGCTGTGATTAACCCCAGCACATTGAGATCGAGTGACGCACTGGCAGAAG           742
         ||||                              |||
Db  744  AATTCATTGAGAAGCCAAGACACATTGAGGTGCAGATGCTCAGATAGAGAACTGGGAATAT   803
```

```
Qy  743  TTGTACACCTTTATGAACGTGACTGCTCACTGCGTCGTGCAGGGTCGTCAGCGTGCAAA     802
         |||                          |||
Db  804  TCTTGCCACTTGTATGAGCGGGCGGCTGCTCCATCCCAGCCGCCAGGAAGTGGTGGTAGAGA  863

Qy  803  TTGGCGCAGCAGCATTTGGATCCAGAACTGCGTGATCGCGTGATCGCGGATGCAGTAA      862
         |||                            |||
Db  864  TTGCCCCTGCTACCCCACCTGCACCTCGGTCACGCGTCACCAGTGACTGCTGTCA         923

Qy  863  AGTTCTGCGCCTCCATTGGTTACCAGGGCGCGGAACCGGTGGAATTCTTGGTCGATGAAA    922
         |||                            |||
Db  924  AACTTGCCAAGCAGGTGGCTATGGAGAATGACGAGGCACTGTGGACACCGTGACTG        983

Qy  923  AGGGCAACCCAGTGTTTCATCGAAATGAACCCACGTTGAAGCGCTGGCACGTGACCTG      982
         |||                            |||
Db  984  ATGGCAAGCACTACTTCATCGAGGTCAATTCCCGCCTGCAGGTGGAGCACACGGTCACTG   1043

Qy  983  AAGAAGTCACCGAGGTGGACCTGGTGAAGGGCGCGTTGGCGTCTGCTGTGTGCAACCT    1042
         |||                            |||
Db 1044  AGGAGATTACAGATGTGGACCTGGTCTCAGATCCGAATCAATGGTTGTGCCATTCAGTGTC  1103

Qy 1043  TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACACGGTCGCAGTGCC        1102
         |||                            |||
Db 1104  TGCCTGACCTAGGCCTGCGGCAGGAAAACATCCGAATCAATGGTTGTGCCATTCAGTGTC   1163

Qy 1103  GCATCACCACGGCAAGATCCAAACAACACGCGTTCCGCCGCAGACTATCACCCGT        1162
         |||                            |||
Db 1164  GGGTCACCACTGAGGACCTGCACGCCAGTTCCGCCAGTTGCGACAATGTTGCCATTGAGGTTT  1223

Qy 1163  ACCGGCTCACCAGGCCGAGCCGGCCTGTG---CAGTCAGCTGCGGTGGCG             1219
         |||                            |||
Db 1224  TCCGGAGTGGTGAGGCATGGCATCCGTGGACAATGCCTCAGCATTCAGGGGAGCTG       1283

Qy 1220  AAATCACCCGCACTTTGACTCGTGATGCTGTGAAATGACCTGCCGTGGTTCCGACTTTG    1279
         |||                            |||
Db 1284  TCATATCCCCACTATGACTCCCTGTCGTCAAGTCATTGCCCATGGCAAGACCACC        1343

Qy 1280  AAACTGCTGTGTCGTCGCACAGCGCGTTGGCTGAGTCTGAGTCTTCGTGTGTTGTCAA     1339
         |||                            |||
Db 1344  CTACAGCTGCCACCAAGATGAGCAGAGCCCTGGCGGGAGTTCCGTGTCCGGAGGTGTAAAGA  1403

Qy 1340  CCAACATTGGTTCTTGCGTTGCTGCCAGAGAGGACTTCACTTCCAAGCGCATCG         1399
         |||                            |||
Db 1404  CCAACATCCCCTTCCTGCAGCAGCACCAGCAGTTCTCTAGCGGCATTCCATTGTGG       1463

Qy 1400  CCACCGGATTTATCGGCGATCACCACCCCTTCAGCGTCCACCTGCGGCGGATGATGAGC    1459
         |||                            |||
Db 1464  ACACCCAGTTGTCATCGATGAGACCCCGAGCTGGTTCCAGCTGCGGCCTGCACAGAACCGGG  1523

Qy 1460  AGGGACGCATCCTGGATTACTTGGACAGATGTCACCGTGAACAAGCCTCATCGTGTGCGTC  1519
         |||                            |||
Db 1524  CCCAGAAGTTGCTACATTACCTTGACACGTCATGGTCAATGCCCTACCACTCCAATCC      1583

Qy 1520  CAAAGGATGTGCAGCAACAATCGATGATAAGC------TGCCCAACATCAAGGATCTGCCAC  1573
         |||
Db 1584  CCGTCAAGGTCAGTCCAGACGCCCTGTGGGACCCCCATTGTTCCTGTGGTGCCCATAGGCCCAC  1643

Qy 1574  TGCCACGCGGTTCCCGTGACCGCCTGAAGCGCTTGGCCTGAGCTTGCCTGCGTTGCATC    1633
         |||
Db 1644  CCCCAGCTGGTTTCAGAGAGACATCCTTCTGCGGACATCCTTCTCCTATATTCTTCAAGGCTG  1703

Qy 1634  TCCGTGAGCAGGACGGACCTGCAGTTACTGACGTACTGATACCACCTTCCGGATGCACCAGTCTT  1693
         |||
Db 1704  TGCGGAATCACCAGGGGCTGCGTCGCTAATGGACACGCTCACCCCACCAGTCAC         1763

Qy 1694  TGCTTGCGACCCGGCGTCCCGTCATTCGCGACGGTTCGCTGTATATTGTCGAAAGC       1753
         |||
Db 1764  TACTTGCCACTAGAGTGCGCACACAGGATCTCAAAAAGATTGCACCCTACGTTGCCCACA     1823

Qy 1754  TGACTCCTGAGCTTTGTCCGTGAGTCGCAGATCCTGGCGATCCGCACTGGCGAAGAAG    1813
         |||
Db 1824  ACTTCAACACCTCTTCAGCATAGAGAACTGGGAGGAGGAGGAGAACTGGGAGGGGCCATGC    1883
```

us-09-974-973-1.rnpb

Wed Apr 9 09:38:22 2003

RESULT 6
US-09-880-107-3029
; Sequence 3029, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 44921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3029
;   LENGTH: 4017
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 S72370
US-09-880-107-3029

Query Match   19.4%;   Score 674.2; DB 10; Length 4017;
Best Local Similarity   54.3%;   Pred. No. 1.6e-198;
Matches 1523; Conservative   0; Mismatches 1258; Indels   24; Gaps   7;

Qy   86 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCAC 145
                |||||| || ||||| | | |||||| || | ||||||| ||| ||| || ||
Db   192 TCAAGAAAGTCATGGTGGCCAACAGAGGTGAGATTGCCATCCGTGCGTGTTCCGGGCCTGCA 251

Qy   146 TCGAAACCGGTGCAGCCACCGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACC 205
                 ||||| ||| || |||| || ||| | ||| |||||| || || ||| || ||
Db   252 CGGAGCTGGGCATCCGACCGTAGCCATCTACTCTGAGCAGGACACGGGCCAGATGCACC 311

Qy   206 GCTCTTTTGCTTCGAAGCTGTCCGCATTGGTACTGGACTGAAGGCTCACCAGTCAAGGCGTACC 265
               ||| | || || || || || || || || | | ||||| || |||| || |||
Db   312 GGCAGAAAGCAGATGAAGCCTATCTCATCGGCCGCGGTCCCGTGCAGGCCTACC 371

Qy   266 TGGACATCGATGAGAAATTATCGGTCAGCTAGAAAAGTAAAGCAGATGCTATTTACCCGG 325
                |||| | | || || || ||| | |||| ||| | || | |||
Db   372 TGCATCCCAGACATCATCAAGGTGGCCAAGGAGAACAACGTAGATGCAGTGCACCCTG 431

Qy   326 GATATGGCTTCCTGTCTGAAAATGCCCAGTTGCCCGGCGAGTGCGCCGGAAAACGGCATTA 385
               | || ||| | |||| ||| | | | ||| || |||| || ||| ||||| ||
Db   432 GCTACGGGTTCCTCTCTGAGCGAGCGGACTTCGCCCAGGCCTGCCAGGATGCAGGGGGTCC 491

Qy   386 CTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTAA 445
               |||| ||| ||| |||| | || || ||| || ||| || | || || ||||
Db   492 GGTTTATTGGTCCAAGCCACCAGAGCTGGTCCGCAAGAGTGTGGAGACAAGGTGGACGCCCGGG 551

Qy   446 CCGCGCGAAGAGAGGCTGGTCTGCCAGTTTGCGG---AATCCACCCGGAGCAAAAACA 502
               | | | |||||||| | ||| ||||| || || | || || || | ||||||
Db   552 CCATCGCCATTGCTGCGGGTGTTCCCTGGTGTTCCCGGCCACAGATGCCCCCATCACGTCCC 611

Qy   503 TCGATGACATCGTTAAAAGCGCTGAAAGGCCAGACTTACCCCATCTTTGTGAAAGGCAGTTG 562
                | || || | ||| | | | || || | ||| | ||| ||| | || || ||
Db   612 TGCATGAGGCCCACGAGTTCTCCAACACTACGGCTTCCCCATCATCTTCAAGGCGGCCT 671

Qy   563 CCGGTGGTGGCGGACGGACCGGTATGCGGTTGTTTCTTCACCTGATGAGCTCCGCAAATTGG 622
               | | ||| | | | | | || || || | |||| || ||| | || |||||
Db   672 ATGGGGTGGGGTGGCATGAGGGTCGTGCACAGCTACGAGGAGCTGGAGGAGAATT 731

Qy   623 CAACAGAAGCCATCTCGTGAAGCTGAAGCGGCATTCGGGCGACGGTTCGGTAATGTCGAAC 682
               | ||||||| ||| ||| || || || ||| ||| | | || || | | ||||
Db   732 ACACCCGGGCCTACTCAGAGCCTTTGGGAATGGGGCGCTGTTTGTGGAGA 791

Qy   683 GTGCTGTGATTAACCCCAGCACATTGAAGTGCAGACACTGGAGATCCTTGGCGATCCCACTGGAGAAG 742
               |||| | || || ||||| | | || | |||||||

Qy   1814 GTTTCCTCTTTGAGGATCCGTGGGACAGAGCTCGACGAGCTGCGCGAGGCGATGCCGAATG 1873
               |||| || ||| || ||| |
Db   1884 GCTTCTTTGTATGAGTGCCCCTGGCCCCTGGCCGGCGGCCTCCAGGAGCTCCGGGAGCTCCCCACA 1943

Qy   1874 TAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGATACACCCGTACCCAGACT 1933
               || || ||| || |||
Db   1944 TCCCATTCCAGATGCTACTGAGGGGCGGGCCAATGCTGTGGCTACACCCTGACA 2003

Qy   1934 CCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGCGTGGACATCTTCCGCATCT 1993
               | | || | | | | || | | || || || || || ||| || ||||
Db   2004 ACGTGGTCTTCAAGTTCTGTGAGGTGGCCAAAGAGAATGGCATGGACGTGCTTCCGGATCT 2063

Qy   1994 TCGACGCGCTTAACGACGTCTCCCAGATGGTCTCAGCAATCGACGCAGTCCTGAGACCA 2053
               |||| | || || || | | || | || ||||| ||| | ||
Db   2064 TTGACTCCCTTAACTACCTGCCAAACATGTGCTGGCATGGAAGCAG---CTGGCAGTG 2120

Qy   2054 ACACCCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCTGATGCAAATGAAA 2113
               | |||| | ||| || | ||||| | | | ||| || || || | ||||
Db   2121 CTGGGGGTGTGGTGGAAGCTGCCATCTCCTACACACGGGTGACGTGGTGCTGCACCCAGTCGCA 2180

Qy   2114 AGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGAGGATCGTCAAGTCTGGCGCTC 2173
               ||| |||| | | | |||| | | | || || | || | || | | |||
Db   2181 CTAAATAACTACTCGGAGTACTACATGGGCTTAGCTGAAGAACTGCTGCGAGCCCGGCACTC 2240

Qy   2174 ACATTCTGGCCATTAAGGATATGGCTTAGCTGCTTCGCCCAGCTGCTGCGGGTAACCAAGCTGG 2233
               | | || ||| | | || |
Db   2241 ACATCCTCTGCATTAAGGACATGGCGCAGGCCTGCGAGGCTGGAAGCCTGCACCCCATGTCGTGG 2300

Qy   2234 TCACCGCACTGCG---CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACAGACACTG 2290
               ||| | ||| | || | | | | || || || | ||
Db   2301 TCAGCTCCCTCCCGGACACCGGTTCCCGGACCTCCCACTGCACATCCCATGCACAT 2360

Qy   2291 CGGGATGGCCAGTTGGCTACCTACTTTGCTCGCAGTCAAGCTGTCACGTTGACG 2350
               | | || |||| || || | ||| | |
Db   2361 CAGGGTCAGGTGGTGCAGCCATGTTGGCCTGTGCAGCAGCTTGCGGCTGGGCTGCTGGTGCTGTG 2420

Qy   2351 GTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCTCCCTGTCTGCCATTCCTGTCTCGCCATTGTTGCTG 2410
               | || | || | || | | || || ||
Db   2421 TGGCAGTCGACTCTATGTCTGGATGACCTCACAGCCCAGCATGGGGCGTGTTCTGAGCCTCGGAGC 2480

Qy   2411 CATTCGCGCACACCCGTCGCGATACCGTGCCGGTTGAGCCTCGTTGTCTTCTGACCTCGAGC 2470
               | || | | | | || || | ||| ||| | || | || | || || ||
Db   2481 GTACCAAAGGAGACTCCTCTGGACACAGAGGTACCCCGTGGGACGTGTGTTTGACTACAGTG 2540

Qy   2471 CGTACTGGGAAGCTGTGTCGCGGACTGTACCTGCCATTTGA------GTCTGGAACCCCAG 2524
               | | | ||| || | | | | | |||| |||| ||| ||| ||||
Db   2541 AGTATTTGGGAAGGGGGCTGTATGCAGCCTTTGATTGCACGGCGGCTACCGGCTACCTACGAAGT 2600

Qy   2525 GCCAACACGGTTCGCGTCTACCGCCACGAAATCCCAGCAGTTGTCCAACCTGCCGTG 2584
               | | | | || | | | | || | | | | ||| | | | ||| ||
Db   2601 CTGGCAACTCAGACGTGTATGAGAATGAGAATCAGAGTGTATCACACCAACCTACACT 2660

Qy   2585 CTGGCAACTCAGACGTGTATGAGAATGAGAATCAGAGTGTATCACACCAACCTACACT 2644
               | | | | || | || | || ||| | | | | | | | || || ||| | ||
Db   2661 CACAGGCCACCCCACTGGGGCCTTGCTTCGATCGCTTGCAGCTCATCGAAGACAACTACGCAG 2720

Qy   2645 CCGGTAATGAGATGCTGGGACGGCCCAAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCG 2704
               || ||| | | ||||| | |||| ||| | || || || ||| ||| ||
Db   2721 AGGCTAACAGATGCTGGGGACCTGATCCATCAAGGTGACACCATCATTGTGGCAGTTGTGGGGGG 2780

Qy   2705 ACCTGGCCACTCCACCTGGTTGTGTCGCGGGTGTAGATCGCAGCAGACTTTGCTGCAGACCCCAC 2764
               |||| | ||| || || ||| | | || || | | || | || || ||
Db   2781 ATCTGGCCCCAGTTCATGGTGCAGAACCGGTTGAGCGGCAGAAGCTCAGGCAG 2840

Qy   2765 AAAAGTACGACATCCCAGACTCTGTCATCGGCCGTTCCTGCGCGGCGGAGCTTGGTAACCCTC 2824
               || ||| ||| || || | | ||| || | | || | | || |||
Db   2841 AAGAGCTGTCCTTCCCCCGCTCTGTGTGGGAGTTCCTCGAGGTCGTACACTTGGCATTCCCC 2900

Qy   2825 CAGGTGGCTGGCCAGAACCACTGCGCACCCGCGACTGGCGAGGCC 2869
               | ||| ||| ||| | | | | || |||
Db   2901 ATGGGGGTTTTCCCTGAACCCTTCCGTTCTAAGGTGCTAACGGACC 2945

```
Db   792  AGTTCATCGAGAAGCCACGGCAGATCTTGGGGACCAGTATGGGAACA  851


Qy   743  TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTGTTGTCGAAA  802
          | ||| ||  |||| ||| || ||||| |||| ||||| |||| ||| |||||||  |||||
Db   852  TCCTGCACCTGTACGACGACTGCTCCATCCAGCGCGGCCACCAGAAGGTGGTCGAGA  911

Qy   803  TTGGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGGTGTCGATGCGAGTAA  862
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   912  TTGCCCCGGCGCCACCTGGACCCGAGCTTCGGACTCGGCCTCACCAGCGACTCTGTGA  971

Qy   863  AGTTCTGCCGCTCCATTGGTTACCAGGGCGGCGGGAATCCGGTGAATTCTTGTGATGAAA  922
          ||||||||| ||||  ||||  |||||||||||||  |||| |||||| |||||| |||||
Db   972  AACTCGCTACAAACAGGTGGGCTACGAGAACCGCAGGCACCGGAGTTCCTGCTGACACAGGC  1031

Qy   923  AGGGCAACCACGTTTTCATCGAAATGAACCCAGTATCCAGGTTGAGCACACCGTGACTG  982
          ||||| |||||||| ||||| ||||||||| || |||| |||||| |||||||| ||||
Db  1032  ACGGCAAGCACTACTTCATGAGTGGTCAACTCTCCCGCCTCCAGGTGGTGGACACGGTCACAG  1091

Qy   983  AAGAAGTCACCGAGGTGGAACCTCTGACCGCGCAGATGCGCATGTCCGTGCTGGTGTCAGTCCT  1042
          |||||  |||||||||| || ||  ||||| ||||||| ||||| |||| |||||| ||  ||
Db  1092  AGGAGATCACCGACGTAGACCTGGTCGCAGGACATCTGTCGATCTGCATCCACCAGTCGCC  1151

Qy  1043  TGAACGGAATTGGGTCTGACCCAAGATGAAGGCGCAGATGCCGCAGACCCACGGTCGCACTGCAGTGCC  1102
          ||||||||||| |||||||||  ||||| || || || ||| |||| ||||  |||| ||||||  ||
Db  1152  TACCCGACCTGGGGCGCAGGAGACCGTAGACGTCGATGCTGGTGCGGTGCCATCCAGTGCC  1211

Qy  1103  GCATCACCACGGAAGATCCAAACAACGGCTTCGGCCCAGATACCGGTGCATCCAGTCGT  1162
          |||||| || ||||  |||| ||  |||||||| |||| || |||| || ||||||
Db  1212  GGGTCACCACGGACACCCGACGTTCCAGCCGGACACCGGCCATTGAGGTGT  1271

Qy  1163  ACCGGTCACCAGGCCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCGGTGGCG  1219
          ||||| ||| |||| ||||| ||| ||||    |||||  ||| |  |  |||||||||
Db  1272  TCCGGAGGGGAGAGGGCATGGCATCGGCTCGGCATAATGCTTCCCGCCTTCCAAGGAGCCG  1331

Qy  1220  AAATCACCGCACACTTGACTCCATGCTGTGTGAAAATGACCTGCCGTTCCGACTTTG  1279
          ||| |||||  ||||||| || || ||  |||| ||||| ||| |||| |||
Db  1332  TCATCTCGCCCCACTACGACTCCCTGCTGGTCAAGTCATTGCCCACGGCAAAGACCACC  1391

Qy  1280  AAACTGTGTTGCTCCGCAGCGCGGCGGCCCTTGGCTGAGTTCACCGTGTCGTGTTGCAA  1339
          |||||| ||| |||||  |||||||||  |||||||| ||| |||||||||||  ||
Db  1392  CCACGGCCGCCAAGAGCAGAGGGCCCTGGAGTTCCGTCCGGAGTTCCAGGTGTGAAGA  1451

Qy  1340  CCAACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGACTTCACTTCCAAGCGCATCG  1399
          ||||| ||| |||  ||  ||||| |||||| ||||||||| ||  || ||||||||
Db  1452  CCAACATGCCGTTCCTGCAGAATGTGCTCAACACAGCAGTTCCTGGCAGGCACTGTGG  1511

Qy  1400  CCACCGGATTTATCGGCGATCACCACCACCTCCTTCAGGCTCCACCTGCGCGGATGATGAGC  1459
          || ||  ||| || ||||  |||||||||| ||  |||| || || |||||||||||
Db  1512  ACACCCAGTTCATCGAGAATCCCAGAACCCAGCGTGTTCCAGCTGCGCGGACAGAACCGGG  1571

Qy  1460  AGGGACGCATCCTCGGATTACTTGGCGAGATGTCACCGTGAA----CAAGCCTCATCGTGTG  1515
          ||||  || |||| ||| |  || || |  || ||||| ||    ||||| |||| |||
Db  1572  CCCAAAAGTCTGTTGCACTACTACTTCGGCCCCATGCGTGTAAACGGTCTTCAACCTGCAC  1631

Qy  1516  CGTCCAAGGATGTGTGCAGCACCAATGCAATAAGCTGCCCAACATCAAGGATCTGCCCACTG  1575
          || || ||| |  |||  || || ||  ||| || ||||| || |||||| || |||
Db  1632  CGGTTCAAGGCCCTGACCCGCAGCAGCTTGCGCCTGCCAGCCGCCGCCATAGGCCCGC  1691

Qy  1576  CCACG------CGGTTTCCGGTACCGCAGCGCCTGGGCCGACTCTTGTCGTCGTCGTGATC  1633
          |||||      ||| ||| || ||  |||| |||| |||| | ||  ||  || |||
Db  1692  CCGGTCAAGGGTTTCAGAGACATCTGCGCGACCCCCTATGTTGCCCACA  1751

Qy  1634  TCGGTGAGGACGCACTGCTCGCAGTTACTGATACACCTTCCGCGATGCCAGCCAGTCTT  1693
          |||| ||| ||||| || || || |||  ||||| |||||| || |  || |||
Db  1752  TGCGGAACCACCCCGGCTGCTGATGATGCCACGACCCTTCAGGGACGCCCCACCAGTCAC  1811

Qy  1694  TGCTTGCGACACCCGAGTCCGCTCATTCGCACTGACTCGCGAGCCGTGCCAAAGC  1753
          ||||| ||  || |||| ||| ||||  ||||||||  |||||  |||| |||
Db  1812  TGCTGCCACTGGTGTGCGACCCCACGTGGTGGTGGTTCCTGCACGGGCTACATCGGTGTCCCCCACA  1871

Qy  1754  TGACTTCCTGAGCTTTTTGTTCCGTGGAGGGCTTGGGGGCGCGACCTACGATGTGGCGATGC  1813
          ||| |||| |||| |||   |||| || ||| || ||||  |||| |||||  ||| |||
Db  1872  ACTTCAGCAAGCTCTTCAGCATGAGAACTCGGGGAGGAGAACTGGGGAGGACTACTGAAGGACC  1931


Qy  1814  GTTTCCTCTTGAGGATCCGTGGGACAGGCTCGCAGCGGCTCGACGAGCTGCGCGAGGCGGATGCCGAATG  1873
          | ||| || |||  ||||||  || |||||| |||| |||| | ||||| |||  ||| |||||||
Db  1932  GCTTCCTGTATGAGTGCCCCGGCCCGTGCGGGGGTGCAGGAGCTCGGGAGCTCATCCCCAACA  1991

Qy  1874  TAAACATTCAGATGCTGTTCGCGCGCGCCGACAACCGTGGGCGCCGCCAATGTGGGATACACCCGTACCAGACT  1933
          ||| ||| || ||| || || ||| | |||| | || |||| |  |  || ||||||| ||  ||||  |
Db  1992  TCCGTTTCCAGATGCTGCTGGCGTGCGGGGGCCAATGTGTGGCTACACCAAGACCACCCAGACA  2051

Qy  1934  CCGTCTGCGCGGCGTTTGTTAAGGAAGCTGCCAGCTGCTGGACATCTTCCGGACTCT  1993
          |  |  | ||||  ||  ||  || || ||| || |||| ||| ||  ||| |||
Db  2052  ACGTGGTCTTCAAGTTCGTGTGAAGTGGCCAAAGAGAATGGCCAAAGAGATGGTCTTCCGTGTGT  2111

Qy  1994  TCGACGCGCCTAAACGACGTCTCCCAGATGAATGACCGTATCCAGGTTGAGCACCGTGACTG  2053
          | ||  |  |  ||| ||| |||| ||||||| |  ||||||||| ||| ||||
Db  2112  TTGACTCCCTCAACTACTTGCCCAACACTTCCCGCCCGCTGCTGGGCCATGGGCATGA---CAGGAAGTG  2168

Qy  2054  ACACCGCGGTAGCCGAGGTGGCTATGCCTATTCTGCTAAAGATGCTACGAGAGAAATGAAA  2113
          ||| |  ||||  ||||| || ||| |  || |  || || |||||||||||| |||||
Db  2169  CCGGAGGCGGTGGTGGCGAGGTGGCTGGCTGCTGGTGGGAGATCGTCAGAGAGAGAGCCGCA  2228

Qy  2114  AGTCTACACCCTGGATTACTACCTAAAGATGGCTATGGCTTGTGTCTTCTGCTTCGCCCCGGT  2173
          |||  |||||| ||| | | ||||  |||||| ||  ||| | || |  ||| ||||
Db  2229  CCAAGTACTCACTGCAGTATCACATACAACGGGCTGGCTCTGAAGGCGTTGGTGCGAGTGGGACCC  2288

Qy  2174  ACATTCTGGCCATTAAGGATGATAGGCTACCCGCCCAGATACCGGAACTATCACCGCGT  2233
          ||||  ||  |||| | |||||| || |   | ||| | |  |||  |||||
Db  2289  ACATCCTGTCATCAAGGACATGGCCGGCGGTGTCTGAAGCCGCTGTCTGAAGCAGCTGCTGG  2348

Qy  2234  TCACCCGACTCG---CCGTGAATTCGATCTGCAGTCACTGTCAGCTCGGTGGCG  2290
          | |||| || |   |  ||| || ||  |  |  ||||  || ||| ||||||
Db  2349  TCAGCTCCCTCCGGGACCGCGTTCCCCGGACTTGCTACTTCCCCCGGACACCGT  2408

Qy  2291  CGGGTGGCCAGTTGGCTACCTACTTTGCTGCCTACTGCCGCAGCGGCAGGTCGGTGGATGATGACG  2350
          |||| |||| ||| ||||  |||  ||  ||||| ||| | ||| ||||| ||||  ||  |||
Db  2409  CAGGGGCAGGCGGCGGTTGCAGCCATGCCCGTGCCTGGGACTGTGGTGAGTGGTGGAGTCCTGGGTGATG  2468

Qy  2351  GTGCTTCCGGACCACTGCTCTGGCACCACCACTCCCACCACCTCCTGTTCTGCTCTG  2410
          || || |||| ||| ||  | ||||||||||| |  |||| || ||   |||
Db  2469  TGGCAGCTGATTCCATGCTGGGGGATGACTTCACAGCCCAGCCCATGGCCGAGGCCT  2528

Qy  2411  CATTCGGCGCACACCCGTCGCGGATACCGGTGTCGTGTTGACCTCGAGC  2470
          || |||| || || ||   ||  || |||||||  | ||| | || |||
Db  2529  GTACCAGAGAGGACTCCCCCTGGGTGTTTGCCCGACGTGTTGACTACAGTG  2588

Qy  2471  CGTACTGGAAGCTGTGCGCGGGACTGTACCGTGCCATTGAGTTCTGAACCCCAGGCCCAA  2530
          |||| |||||||| |||| || |||| ||  |||| | ||| |  || |||| || |||
Db  2589  AGTACTGGGAGGGGCTCGGGAGCTCGCTGTACGGCGCCCTGCACCACTGATGAAGT  2648

Qy  2531  CCGG------TCGGTCTACCGCGGTCATCGTGTG  2584
          ||||      ||||  || |||| |||  ||||||
Db  2649  CTGGCAACTCGGACGTGTATGAAAATGAGATCCCACGCTGGCCCAGTACACCAACCTGCACT  2708

Qy  2585  CACAGGCCACCACCGCTGGGCCTTGCCTGCTCTGATCCTCGAGCTCGAAGACAACTACCAG  2644
          || ||||| ||  | || ||  ||  || ||||| ||  || ||| ||||| |  ||||||
Db  2709  TCCAGCCCCAACATGGGCCTTGGGCTTCGGGCTCCCATAGGAGGTTCAAGGAGGTCATGTGG  2768

Qy  2645  CCGTTAATGATGATGGTTGGTGGTGGCGGCCAACCATCCTTCCAAGGTTGTTGGCG  2704
          ||||  ||||||||  || ||||  |||||||||| ||  || | ||||| |||||
Db  2769  AGGCCAACCAAGATGTGGGGCGATCTCATCAAGGTGACGCCCCTCCTCCAAGATCGTGGGGG  2828

Qy  2705  ACCTCGCACTCCACCTGTTGGTGGCGGTGTAGATCCAGCAGAGACTTTGCTGCAGACCCAC  2764
          ||||| || ||||| || |||| ||| ||| || |||| |||||| |||| || || |||
Db  2829  ACCTGGGCCCCAGTTTATGTGCAGAATGGATTGAGCCGGCCCGAGAGCCGAAGTCAGGCGG  2888

Qy  2765  AAAAGTACGACATCCCAGACTCTGTCGTCTGGGCCTTGTCATCGGCCGTTCGGTAACCCTC  2824
          || | ||| ||| | |||  |  ||  |||| |||||  |||  |||| || ||
Db  2889  AAGAGCTGTCTTTCCCCGGTCCGGTGGTTCCTGGAGTTCCTGCACGGGCTACATCCGGTGTCCCCC  2948

Qy  2825  CAGGTGCTGCCAGAACCACTGCCGACCCCGCCACTGCGAAGGCC  2869

Db  2949  ATGGGGGGTTTCCCCGAACCCTTTCGGCTTCTAAGGTACTGAAGGACC  2993
```

RESULT 7

US-09-815-242-6709

; Sequence 6709, Application US/0981524 2

; Patent No. US2002006156 9A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6709
; LENGTH: 3429
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3429)

US-09-815-242-6709

Query Match   14.5%; Score 502.8; DB 10; Length 3429;
Best Local Similarity 49.0%; Pred. No. 2.7e-145;
Matches 1657; Conservative   0; Mismatches 1672; Indels   51; Gaps   10;

Qy   86 TCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCAC 145
          | |||||||| |||||| |||||| |||||||| ||||||||| |||| ||||||||
Db   2 TGAAAAAGGTATTAGTTGCCAATCGTGGCGAAATTGCGATTCGAATTTTTAGAGCTTGTA   61

Qy   146 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGAGGTCATCCCACC 205
          || ||| |||| ||||||| ||| ||||||| ||| |||| |||| |||||||| ||||
Db   62 CAGAATTAGATATCCGTACAGAGCGATTTATGCTGCAGAAGATGAGTATTCTGTTCATC 121

Qy   206 GCTCTTTTGCTTCTGAAGCTGCCATTGGTACTGGTAAGGCTCACCAGTCAAGGCGTACC 265
          | || |||||||| |||| ||||||||||| |||||||| |||| ||||||||||||
Db   122 GTTTCAAAGCAGATGAAGCGTATTTAGTTGGTAAGGAGAAAAAACCAATTGAAGCTATT 181

Qy   266 TGGACATCGCGATGAAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCCGG 325
          ||||| | || ||||| |||| || |||||||||||| |||| ||||||||||||
Db   182 TAGACATTGAAAATATTATCCAAATTGCCAAAAAATCGGAGCAGAGCCATTCATCCTG 241

Qy   326 GATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGCGCGGGAAAACGGCATTA 385
          ||| | ||||||| ||| || |||||||| | || ||||||| ||| || |||||||
Db   242 GTTATGGCTTTTTTATCAGAGAACTTACGCTTCGCTTCGCAACGTTGTGAAGAAGAAATCA 301

Qy   386 CTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATCGTGATAGTCTCGTGCGGTAA 445
          | || ||||| ||||| ||||||| ||||||||| |||||||| |||| |||||| |||
Db   302 TTTTTGTTCGGACCTAAAACACATCATTTAGATATTTTTGGCGGATAAAATTAAAGCGAAAG 361

Qy   446 CCGCCGCAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATC---CACCCCGAGCAAAAACA 502
          | |||||| | ||||||| ||||| |||||||||||| ||||||||||| ||| |||
Db   362 AAGCAGCTGTAGCTGCGGGGATTGCTTCGATTCCAGGCTCAGACGGCCCAGTAGCAACGG 421

Qy   503 TCGATGACATCGTAAAAGCGTAAACGCTGGAGGCCAGACTTACCCATCTTTGTAAGGCAGTTG 562
          || ||||||||||| ||||| || || |||||| ||| ||||||| | || |||||||||
Db   422 TGGAAGAGGTCGTAGCTTTTGGTGAAACGCATGGCTTTCTATCATGATTAAAGCTGCTT 481

Qy   563 CCGGTGGTGGCGGACCGCGGTATGCGCTTTGTTTCTTCACCTGATGAGCTCCGCAAATTGG 622
          ||||||||||| ||||||||| ||| ||||| ||||||||||||| |||| ||||| |||
Db   482 TAGGTGGCGGCGGTCGGGGATGCGGCCTTGCCACCATGGCCAAAGAGCACAGAGAAGGTT 541

Qy   623 CAACAGAGCATCTCGTGAAGCTGAGCGCCATTCGGCGACGGTTCGGTATATGTCGAAC 682
          | ||| ||||| |||||||||||||||||| | ||||| |||||||| ||| |||||
Db   542 ACGAGAGCCAAAAGTGAAGCGAAAGCAGCCTTTGGTTCTGACGAGGTTTATGTTGAAA 601

Qy   683 GTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGGCACTGGAGAAG 742
          |||| ||||| ||||| |||| |||| |||| ||||||| |||||| ||||| |||||
Db   602 AGTATATTTCTAATCCTAAACATATCGAAGTACAAATTTAGGCGATCATCATGGGAATG 661

Qy   743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCCGTCGTCACCAAAAGTTGTCCAAA 802
          ||| || |||| | |||||||| || || ||||||| ||||| ||| ||||| |||||
Db   662 TCTTGCATTTATTTGAACGTGATTGTTTCCGTTCAACGACGCCACCAAAAAGTGGTAGAAG 721

Qy   803 TTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGGATCGCATTGCGGATGCAGTAA 862
          || || || || || |||| ||| ||||| | |||||||| || |||| ||| |||||
Db   722 TCGCACCATGTGTATCAATGAATGAAGAACAACGAGCAGCCATTTGTTCGGCTGCTGTGC 781

Qy   863 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA 922
          |||| |
Db   782 AGTTAATGGCGATGTCGGCTACGTGAATGCGGGAACGGTTGAGTTTTTAGT---AGAAG   838

Qy   923 AGGGCAACCACGTTTCATCGAAATGAACCCACGTATCGAAATGTGACACACCGTGACTG 982
          || || |||| || ||| || |||||||||| |||| || |||||| |||||||||
Db   839 GGGATCAGTTTTATTTATTGAAGTGAATGAATCCTGCGTGTTCAGTGAGTAGAACATATCACAG 898

Qy   983 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACCT 1042
          | ||| | | ||| || || | |||| ||||| | | |||| |||||| ||||||||
Db   899 AAATGATTACAGATATTGATATTGTGATCTGCAATTACAAATTGCACAAGGGCTTGATT   958

Qy   1043 TGAAGGA-----TTGGGTCTGACCCAAGATAAGATCAAGACCACACCGTGCAGCACTGC 1096
          || || | |||||| | | |||| || ||||| ||| | ||||||||| |||
Db   959 TGCATAAAGATATGCATTTGCCAAAACAAAACAATTGACATTAAAAGCGCGGCTATTC 1018

Qy   1097 AGTGCCGCATCACCACGGAAGATCCACCACAACGCGTTCCGCCCGATACCGGAACTATCA 1156
          | ||| |||| | || || | || ||| || || |||| | ||| |||| |||| |||
Db   1019 AATGTCGGATTACGACTGAAGATCCCCTGAACCAGTTTATGCCAGATACAGGGAAAATCG 1078

Qy   1157 CCGCGTACCGCTCACCAGGCGAGCTCGGGTTGGTGTTGGTTCCGTTAGTGGCAGAGAAATGCTTATTCTG 1213
          | || ||||| | ||| ||| ||| || | ||| | |||||| || ---|||||| || ||||| |
Db   1079 ATACGTATCGTTCACCAGTGGCGTTGGTGTGGGTGCCTTGTACTCATGGCTGGCATGATGCCGTGCG 1138

Qy   1214 GTGGCGAAATCACCGCACTTTGACTCCATGCTGAAAGTCGGGGATCATTCCACC 1273
          ||| | || | |||| ||| ||| ||| | | | |||||| |||||| |||||
Db   1139 GCTACGCTGTGAGCGCCTTACTTTGATTCTTTATTGTTTATTGGTTAAAGTCTGTACGCATGGTTTCT 1198

Qy   1274 ACTTTGAAACTGCTGTTGCTGCACAGCGCGCGTGGCGTGAGTTGCAGTCGTGGTG 1333
          | |||||||| || || |||||| | | |||||| |||||| | ||| ||| ||
Db   1199 CTTTTGAACAGCAATCAGTAAAATGCAACGGTCCTTAAAAGAATTTCGGATTCGTGCG 1258

Qy   1334 TTGCAACCAACATTGGTTTCTTCTTGCGTGCCGTTGCTGCGGGGAAGAGGACTTCACTTCCAAGC 1393
          || ||||| | | || || ||| | |||| | |||| ||| ||||||||| | || |||||
Db   1259 TGAAAACAAATATTCCGTTTTTACAAAATGTTGTGAGCTATCCAGCGTTCAATCTGGCG 1318

Qy   1394 GCATCGCACCGGATTATCGGCGATCACCACCTCCTTCAGGCTCGCCGGATG 1453
          |||| || | || || ||| |||||| || || ||| || ||| || |||
Db   1319 AAGCCAAAACAACCCTTTATTGAAATATACACCTGAATTATTTGAATTTCCTCGTATGCCGCG 1378

Qy   1454 ATGAGCAGGGACGGACCGCATCCTGGATTACTTGGCAGATGTCACCGTGAACATGGGTG 1513
          || | | || |||| | ||| | | |||||| | || || |||| |||||| ||||||
Db   1379 ATCGTGGCAATAAACGATGAAATAAAAGGTAACGGTCAATGCATGGGTTCCCTGGCA 1438

Qy   1514 TGCGTCCAAGG------ATGTTGCAGCACCAATCGATAAGCTGCCAACATCAAGGATC 1567
          | || |||||| |||| |||||| |||| ||| ||||| || |||| || ||
Db   1438

```
Qy  2648  TTAATGAGATGCTGGGACGCGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCCACC  2707
          |||  || |||| ||| || |  ||         |||  ||||  ||  |||| || | |
Db  2576  TGAATTTGATGTTTGGTGATATTGTAAAAGTGACGGCCATCTTCAAAAGTCGTGGGAGACA  2635

Qy  2708  TCGCACTCCACCTGGTTGGTGCGGGTTGAGTCTGACACTTTGTCTGCAGACCCACCACAAA  2767
          | |||||| | ||||||    |  ||||||| ||| || |  | |  | | ||| |||| |
Db  2636  TGGCACTCTTCATGGTTCAAATCATCTGACAGAACAAGATGTTTATGCGGCGTGGGTGAAG  2695

Qy  2768  AGTACGACATCCCAGACTCGTCATCGCGGTTCCTGCGCGGCGGAGCTTGGTAACCCTCCAG  2827
           | || | | ||  | || |   |||  || | |  |  || |  | |||| ||| || ||
Db  2696  AACTAAGTTTCCTGAATCAGTTGTGACTTTTTCCAAGGTGATTTAGGTCAACCAGTTG     2755

Qy  2828  GTGGCTGGCCAGAACAACTACTGCGCGCACCCGGCCGCCTCCGAAGGCAAGGCAC        2887
          ||||  | ||| ||| | |  |  |   ||||  |  | |  |||| | || |
Db  2756  GTGGGGTTCCCGAAAGAGCTCCAACGAATTATTCTAAAGGTCGACCAGCCTTCACAGAGC  2815

Qy  2888  CTCTGACGGAAGTTCCTGAGGAAGAGA----GGCGCACCTCAACCGCGCTGCTGTTCAAG  2943
          || | || |||  | | ||| | ||||    |  | |||    |  | ||   ||  |||
Db  2816  GGCCAGTGACTTAGCAGCACTGGTTGTGAAGTACAAGAGAATTAGCCGAAA          2875

Qy  2944  GAACGTGCAACAGCCTCAACCGCCTGCTGTTC------GGCGCACCTCAACCGCGCTGCGTCTGTG 3046
          | || | |||| || | | ||  | |||                ||  || | |
Db  2876  AAATTGGGTATCAACCTAAATTGGAAGAAGTTTTAAGTTATTTAATGTATCACAAGTAT  2935

Qy  2987  CCGAAGAGTTCTTCCTGAGCACCGCGCGCCGTTCGGCAACACCTCTGCGCGTGACGATCGTG  3046
           | ||||||| || |||  | ||  | ||   | | ||| |   ||  |  | ||| || | |
Db  2936  TTTTAGAGTATCGTCAAAAATACGAAACCTTTGGAGATATAAACCTTATTAGATACGCCAA  2995

Qy  3047  AATTCTTCTACGGACTGGTCGTGAGGGCGCCGGAGACTTTGATCCGCCTGCCAGATGTGCGCA  3106
           | ||||| | ||  |  | | || ||    ||| ||| |  || ||    | ||| | | |
Db  2996  CATTCTTTAATGGTGTATTCGTCAAGGA---GAGACATTGGAAGTTCAAATTGAACGAGGAA  3052

Qy  3107  CCCCACTGCTTGTTGTCGCCTGATGCGGATCTCGAGCCAGACGATAAGGGTATGCGCAATG  3166
              | |||| ||| ||   | | || |  ||| | ||  | || | ||| |||| | | |||
Db  3053  AAACGTTAATTATTCGTTTAGATGAAAATTGCGGCGAACCTGATATTGACGGAAACCGTGTCT  3112

Qy  3167  TTGTGGCCAACGTCAACGGCCAGATCCGCCGCCCAACAGGGCCATGTTGCTGCCACCAT  3226
          |||| |||||  |  |  ||||| |     | |||||||||  || | || | | ||||
Db  3113  TGTTCTTCAATTGAATGGGCAACAGGCAGATTCCTCCACAACAGGGAACAAATCGGCGCAACGA 3172

Qy  3227  CTGTCACCGCGCAACCGCAGAAAAGGCAGATTCCTGAAGGTAGCTCAAGGATG         3286
          |||| || |  | || | ||||| |||   |     || ||| | |  |  |||
Db  3173  CTGCTGTTCAAGTAAAACAAAAGGCGGAACAACTAATAAACGTGGCGATTGGTCAAACGTTGTCT 3232

Qy  3287  TCGCTGG---TGTTGTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGAGATG   3343
          | |||||   |||| || | | |  ||| ||  |||  ||| | || |||| | |  | |
Db  3233  TGTCTGGTTCGTTTCACAAGTATTGGTCAAACGTGGCGATAATTGGTCAAAAGGTGAGAAAAGGTCAGC 3292

Qy  3344  CAGTGCAATCATCGAGGCTATGAAGACAACAATCACTGCTTCGTTGACGGCA          3403
          ||||||||| ||  || | ||||   | |||| | || | |    ||| | |
Db  3293  CGTTGCTGATTACAGAGCCATGAAAATGAAACGACTATCGAAGCAGCGTTTTGCGGGCA   3352

Qy  3404  AGATTGAACGCGTTGTGGTT  3423
          || ||||||| | | | |||
Db  3353  CGGTTGATCACATTTATGTT  3372
```

```
Db  1439  TTGAACGAACGGAGAAAAATATTTTGAAGGCCCCGCGGTGCCGACAGATATTGAAGTTC  1498
          || |||   |||  ||   |||   ||| |||  |  ||| ||| | || | | || |
Qy  1568  TGCCACTGCCACGCGGTTCCCGTGACGCGCCCTGAAGCAGCTTGGCCCTCCAGCCGCGTCATCG 1627

Db  1499  CAGAAAAAGTGATCACAGCTAAAAATATTCTAGACGCTCAAGGAGCAACCGCCGGTCATCG  1558
          | | | |||| |||| |||  ||| ||   |  |||| |  ||   ||| |  || | | |
Qy  1628  GTGATCTCCGTGAGCAGGAGGACGCACTGGCAGTTACTGATACACACCTTCCGCGATGCCACC  1687

Db  1559  ATTGGGTTAAGAATCAAGAAGAGTGTGTTAAATACAGATACCACGTTCCGTGATGCCGCACC  1618
          |||||| ||||||| ||| | | | | |||| ||||| || | |||   | || | | | ||
Qy  1688  AGTCTTGCTTGCTGACCCGAGTCCGTCATTCGCACTGAGCGCTGCGGAGAGGCCGTCG  1747

Db  1619  AAAGTTTATTAGCTACTCGTGTGACAACAACAAGATTTTAAAGCAATTGCTGGACTAACTG  1678
          ||||||||||||| |||||||||||| || | | | ||  ||| | ||  |  || | |||
Qy  1748  CAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGAGGCGGCGGCCACCACGATGTGG  1807

Db  1679  ATGTAGCCCTGGCCTGAGCTGTTTCTAGTGAAATGTGGGTGCGGGTGCGCCTACGTTCGATGGTTG 1738
          ||||| ||| |  || | || | | | |||||||  | | | ||| | | | || || | |
Qy  1808  CGATGCGTTTTCCTCTTTGACGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGC  1867

Db  1739  CCTATCGCTTCTTAACCGAAGACCCATGGCGCACGTTAACTCGTGTCAGTTGATGC  1798
          | ||| ||||||| |||| || ||| ||| | |||||| | ||||||||||||| |
Qy  1748  CAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGAGGCGGCGGCCACCACGATGTGG  1807

Db  1799  CAAACATAACTTTTCAACATGCTGTCAGAGGATCTAATGCGGAGGTAGGCTATCAAATTACC  1858
          | || |||| ||||| | || ||||| |||| |||||||||| | ||||||||||||| |||
Qy  1808  CGATGCGTTTTCCTCTTTGACGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGC  1867

Db  1859  CTGATAATGTTATTGGAAGAATTCATTAAAGAATCTGCTCGCGCAAGGAGTCGATGTTTCC  1918
          |||| || ||| ||||| |||||||| ||||||||||| ||||| | ||||||||||||||
Qy  1868  CGAATGTAAAACATTCAGATGTCGTGCTTCGCGGCCGCAACACCGGTGGGATACACCCGTACC  1927

Db  1919  GAATTTTGATAGCTTAAACTGGATTCCTCAAATGGAAAAAGTATTCAAGTCGTTCGGG  1978
          ||||||||||| | || || |||| ||| |||||||||||| |||||||||||||||
Qy  1988  GCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCTCCAGCAATGACGCAGTCCTGG  2047

Db  1979  ATACCGGAA---AAATGCGGGAAGCACAGCAATTTGTTATACTGGGGACATCAATGATCCAG 2035
          | ||||| |   ||||||||||| || ||||||| |||| | |||||||| |||| |||||
Qy  1988  GCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCTCCAGCAATGACGCAGTCCTGG  2047

Db  2036  CCCGAGCAAATAATAATGTTCAATTTAACCTTGTTATATGGCTGGGTAAAGATATGGAAAAATTTAG 2095
          |||||| ||| || |||||| |||||||| ||||||| ||||||| ||||| ||| ||| |||||
Qy  2048  AGACCAACAACACCGGGTAGCCGGAGTGCTATGCTTATTCTGTGTATTCTCTGATCCAA  2107

Db  2096  GGGCACAAATCATTGCGATTAAAGATATGGCTGGCTTATTGAAACCACCACAAGCTGCTTATC 2155
          ||| || |||||||| | ||||||||||||||||| ||||| |||||||||||| ||| ||
Qy  2168  GCGCTCACATTCTGCCCATTGAATTCGGGATATGGCTGTCGTCTGCCCAGCTGCGGTAACCA  2227

Db  2156  GTTTAATTAGTGAATTAAAGGCAGCACCACGGATTTACCCATTCATCTCCATACTCACGATA  2215
          |||||||||| ||||| |||||||||||| || ||| |||||| |||||||||||||||||
Qy  2156  GTTTAATTAGTGAATTAAAGGCAGCACCACGGATTTACCCATTCATCTCCATACTCACGATA  2215

Db  2216  CTAGTGCAATGGGAATCATCATTATTCAGCAGTAGCAGGCGTTGATATTGTTG  2275
          || |||||||||||||||||||| |||||||||||||||| ||||| |||| |
Qy  2228  AGCTCGGTCACCGACACTGGCTCACCTTTGCTGCAGTCACGTGCACGTGCACACCCACGACA  2287

Db  2276  ACGTCGCAATGAGTGCGATGAGTCGATAGCCTTCAGCTCTGCTTCAGAAGAGAACAACTACGCACACAG 2335
          || |||||||||||||||||||| |||| |||||| |||||||||||||||||||||||||||||||||
Qy  2288  CTGCGGGTGGCCAGTTGCCACTGGCTACCTTTGCGCACTTCACCCTTCATCAGCATGGGA  2347

Db  2336  ATGCTTTAGTGACTGGTCAATGGTGAACGAGCACGCCAACTATTAACATTGATAATGCACAGAAAAATCA 2395
          |||| |||||||||||| |||| |||||| || | ||||||| ||||||||| |||||||||||||||||
Qy  2348  ACGGTGCTTCCGACCACTGTGCTGGCACCACCCGTTCCCCAGCACCCTGCCGTTTCTGACCTCG  2407

Db  2396  ATCATTATTGGGAGAGATGTGCGCATGCATGTATTATCAACCATTGAAAATGGGTTAAATGCAC 2455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2408  CTGCATTCGCCGGCCACACCGGGCCGATACCGGTTGAGCCTCGAGCCTGTTTCTGACCTCG  2467

Db  2456  CGCAACAGAGTCTATATGCATGAAATGCCTGGTGTCAATACTCCTATTCCTACACGCAGC  2515
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2468  AGCCGTACTGGAAGCTGTGCGCGACGTGTACCTGCCCATTTGAGTCTGCAGACCCCAGGCC  2527

Db  2516  AAGCAAAAGCGGTAGGCTTAGGGCACCGTTAGGGGATGAAATCAAAAAAATGTATCACACAG  2575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2528  CAACCCGGTCGCGTCTACCGCCACAGCCCAGCGGACAGTGTTCCAACCTGCCGTGCAC  2587

Qy  2588  AGGCCACCGACTGGCCTTCGATCCGTTCGCTTCGAGACTACACGGCCGCCG  2647
Db  2516  AAGCAAAAGCGGTAGGCTTAGGGCACCGTTAGGGGATGAAATCAAAAAAATGTATCACACAG 2575
```

```
COMPUTER READABLE FORM:
  MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
  COMPUTER: HP Vectra 486/33
  OPERATING SYSTEM: MSDOS version 6.2
  SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/781,986A
  FILING DATE:
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
ATTORNEY/AGENT INFORMATION:
  NAME: Benson, Bob
  REGISTRATION NUMBER: 30,446
  REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (301) 309-8504
  TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 324:
SEQUENCE CHARACTERISTICS:
  LENGTH: 5030 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
US-08-781-986A-324
```

```
Query Match         14.1%;     Score 488.4;  DB 7;   Length 5030;
Best Local Similarity  48.5%;    Pred. No. 1e-140;
Matches 1665;  Conservative    0;  Mismatches 1711;  Indels   60;   Gaps   9;
```

```
Qy      84  ATTCAAAAAGATCTTGGTAGTAGCAAATCGCGGCAAATCGCGGTCCGTGCTTTCCGTGCAGC 143
            ||||| |||| | |||| || |  |||| ||  || ||||
Db    3846  AATAAAAAGTTACTTGTTGCTAACCGTGAGAAAATTGCAATTCGTATATTCAGAGCGGC 3787

Qy     144  ACTCGAAACCGGTGCAGCCACGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCA 203
            |   ||||||||||
Db    3786  GGCAGAATTAGACATCAGCACAGTTGCAATTTATTCGAATGAAGACAAAAGTTCATTACA 3727

Qy     204  CCGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGGCGTA 263
            |||||
Db    3726  TAGATATAAAGCAGATGAATCCTATTAGTTAGTTGGAAGTGATTTAGGTCCTGCTGAAAGTTA 3667

Qy     264  CCTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCC 323
            ||  |||
Db    3666  TTTAAATATTGAGCGTATCATTGATGTAGCAAAACAAGCGAATGTGGATGCGATTCATCC 3607

Qy     324  GGGATATGGCTTCCTGTCTGAAAAATGCCAGCTTGCCCCGCGAGTGCGCGGAAAACGGCAT 383
            |  |||||| |||||||||
Db    3606  TGGCTATGGATTTTAAGTGAAAATGAACAAATTGCGCGTCGTTGTGCAGAAGAGGAAT 3547

Qy     384  TACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGT 443
            ||||||
Db    3546  TAAATTTATTGGTCCTCATTTAGAAACATTTGAGATCGATTTTGAGATAAAGTTAAAGCTCG 3487

Qy     444  AACCGCGCGAGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCACCCCGAGCAAAAAACAT 503
            ||||
Db    3486  TACAACGGCTATCAAGGCAGATTTACCAGTTATTCCTGGTACAGACGGTCCAATTAAATC 3427

Qy     504  CGATGACATCGTTAAAAGCGCTGAAGGCCAGAC---TTACCCCATCTTTGTAAAGGCAGT 560
            |||||
Db    3426  ATATGAATTAGCAAAAGAAATTGCAGAAGAAGCGTGGTTTCCCGCTAATGATTAAAGCCAC 3367

Qy     561  TGCCGGTGGTGGTGGGCGGCGGGTAAAGCGCTATGCGGTGATGCAAATCGTTCGTGAAGAAAT 620
            |||||
Db    3366  AAGTGGTGGCGCGGCGGGTAAAGGTATGAAGGCTATGAAATTGAATTAGAAGATGC 3307

Qy     621  GGCAACAGAAGCGATCTCGTGAAGCTGAAGCTGAAGGCAATTCGGCAGCGCGCATTCGGGC 680
            |||||
Db    3306  TTTCCATAGAGCAAAATCAGAAGCTGAAAAATCATTTGGTAATAGTGAAGTTTACATAGA 3247

Qy     681  ACGTGCTGTGATTAACCCCCAGCACACATTGAAGTGCAGATCCTTGGCGATCGCACTGCAGA 740
            || || |||  ||||| |||||||||
Db    2172  ATCCAAAACAGCGGACGTATTTAAAGATGGTTTCTCACTAGAAATGTGGGCGGTCGTAC 2113
```

```
Db    3246  AAGATACATTGATAATCCAAAGCATATTGAAGTACAAGTCATAGGTGACGAACATGGAAA 3187
            ||||| ||| || | |
Qy     741  AGTTGTACACCTTTATGAACGTGACTGCTCACTGCACCGTCGTGCATCCTGCAAAAAGTTGTCGA 800
            ||| || || | |
Db    3186  TATCGTACACTTATTTGAACGGTGATTGTTCAGTACAACGTCGTCATCAAAAAGTTGTAGA 3127

Qy     801  AATTCGGCCAGCACAGCATTTGATCCAGAACTGCGTGATCGGCATTTGTCGCGATGCAGT 860
            || | ||
Db    3126  AGTTGCACCATCAGTTGGATTATCATCAACATTACGTCAACGTATTTGTGATGCTGCAAT 3067

Qy     861  AAAGTTCTGCCGCTCCATTGGTTACCAGGGGCGCGGGAACCGTGGAATTCTTGGTCGATGA 920
            |||  ||
Db    3066  TCAATTGATGAGAAAATATTAAATATGTCAATGCAGTGTACTGTTGAATTTCTAGTATCTG- 3008

Qy     921  AAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGAC 980
            ||| || |
Db    3007  --GTGACGAATTCTTCTTTATAGAAGTTAACCCTCGTGTACAAGTGCGCTTGGCTCGTGCAAC 2950

Qy     981  TGAAGAAGTCACCGAGGTGGACCTGGTGGAAGCGCAGATGCGCTTGGCTGCTCGTGCAAC 1040
            |||| || || | ||
Db    2949  AGAGATGGTAACAGGAATTGATATTGTTAAGACACAAATTTAGTTGCAGCAGGTGCCGA 2890

Qy    1041  CTTGAAGGAATTCG--------GTCTGACCCAAGATAAGATCAAGACCCACCGTGCAGC 1091
            |||| || |||||
Db    2889  TTTATTTGGTGAAGAGATTAAATATGCCGCAACAAAAAGATATTACAACATTAGGCTATGC 2830

Qy    1092  ACTGCAGTGCCGCATCACCACCGGAGAATCCAAACAACGGCTTCCGCCCAGATACCGGAAC 1151
            || || || |
Db    2829  CATCCCAATGTCGTATTACAACACAGAAGATCCGTTAAATGATTTCGTCTTGACGTGCAGCTCAGCT 2770

Qy    1152  TATCACCGGCGTACCGCTCACCAGGCGACGTGGCGTTTGGTGTACGTCTTGATGCTGGATGGTTT 1211
            || ||||| || |
Db    2769  AATCATTGCTTATCGTTCAAGCGGGGCGTTTGGTGTACGTCTTGATGCTGGATGGTTT 2710

Qy    1212  C---GGTGGCGAAATCACCCGCACACTTGACTCCCATGCTGGTGAAAATGACCTGCCGTCG 1268
            | |
Db    2709  CCAAGGTGCTGAGATATCACCTTATTATGTCATTCATCGTAAAATTATCACACACGC 2650

Qy    1269  TTCCGACTTTGAAACTGCTGCGTGTCGTCGTGCCACGCGCGCGCTGGCTGAGTTCACCGTGTC 1328
            || ||
Db    2649  GATATCATTTAAACAAGCAGAAGAAAAATGGTACGGCTCATTACGGAAGAAGAGGACTTCACTTC 2590

Qy    1329  TGGTTGTTGCAACCAACATTGGTTTCTTGCGTGCCGTTGCTGCGCGGAAGAATAAAAAGTTCACAAG 1388
            ||||||
Db    2589  TGGTGTTAAAACTAATATTCCATTCTTAATTAATGTAATTCTTAATTAATGTAATTCAGGGTGCAACACCTCCTTCAGGCTCCACCTGC 2530

Qy    1389  CAAGCGCATCGCCACCGGATTTATCGGCGATTTATTGAACGAAACACCAGAACTTTTCGACATTCAGCCGTC 1448
            ||
Db    2529  TGGTGATTACACAACTAAATTTATTGAGAAACTGCTTCGGCAGATGTCGAACAAGCCTCA 2470

Qy    1449  GGATGATGAGCAGGGACGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCA 1508
            || ||
Db    2469  TCTAGATAGAGGTACTAAAACATTAGAATATATAGGAATGTGATAATGTAACAATTAATGGTTTCCC 2410

Qy    1509  TGGTGT---------GCGTCCAAAGGATGTTGCAGACCAATCGATAAGCTGCCCCAA 1556
            | |
Db    2409  AAATGTTGAGAAAACGTCGAAACCAGACTATGAATTAGCATCAATTCCAACTGTATCTTC 2350

Qy    1557  CATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCTGAGCAGCTTGGCCCAGC 1616
            |||||
Db    2349  AAGTAAAATCGCTTCATTT--AGTGGTACGAAACCAATTGCTTGATTGAAGTAGGTCCAAA 2293

Qy    1617  CGCGTTTGCTTCGTGATCTCCGTGAGCAGAGCGCACTGGCAGTTACTGATACCACCTTCCG 1676
            |||||
Db    2292  AGGTGTAGCTGAATGGGTTAAAAGCAGAGAAATGCATGGTCTTAACTAACAGATACAACCTTTAG 2233

Qy    1677  CGATGCACACCAGTCTTTTGCTTGCGACCCGAGTCCCGTCATTCGCACTGCACTGAAAGCCTGCGGGC 1736
            ||||| |
Db    2232  AGATGCACACCACCAATCATTATTAGCTACACGAGTTAGAACTAAGGATATGATTAATATCGC 2173

Qy    1737  AGAGGCCGTCGCCAAAGCTGCTGACTCCTGAGCTTTTGTTCGTGAGGCGGCGGCGGCGAC 1796
            |||||
Db    2113
```

```
Qy  1797  CTACGATGTGCGATGCGTTCCTCTTGAGGATCGTGGACAGGCTCGACGAGCTGCG  1856
                ||||| ||  |||  |   ||    |||||  |    || |||   ||||| ||
Db  2112  ATTTGGATGTGGCATATAATTTCTTGAAGGAAAAACCCATGGGAACGACTTGAACGTCTACG  2053

Qy  1857  CGAGGGCGGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGCGCCCAACACCGTGGGATA  1916
                | || | ||||  |||||| |  |||  ||  | ||  ||  | | ||| |||  ||| |
Db  2052  TAAAGCTATTCCAAATGTATTATTGTATTACGTGCTTCAAACGCAGTTGCTTGGTTA  1993

Qy  1917  CACCCCGTACCCAGACTCCGTCTGCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGT  1976
                || |||  |||  || | |   |     | |  || ||| |   |||| |  ||
Db  1992  TAAAAACTATCCTGATAATGTTATTCATAAATTCGTACAAGAAGTGCTAAAGCAGGCAT  1933

Qy  1977  GGATGATCGTGAATTCTTCTACGGACTGGTCGAGGGCCGCGGAGACTTTGATCCGCCTGCC  2036
                | ||||  |  | |  | ||  ||||| || |     ||   |   |   |  ||
Db  1932  AGATGTCTTTAGAATTTTCGATTCATTAAACTGGGTAGATCAAATTGCCAATGA  1873

Qy  2037  CGCAGTTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCGAGCCAGACGATAAGGG  2096
                |  || |  ||  |  || |   ||| | | ||  |    | ||  |||| | | |
Db  1872  AGCAGTACAAGAAGCGGGCAAAATCTCAGAAGGTACTACATTTGTTATACAGTGACATTTT  1813

Qy  2097  CTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGAT  2156
                | |||  ||    || ||| |  || ||| || |  || |   |  || |||  ||
Db  1812  AAATCCTGAGCGATCAAACATTTATATCTTTAGAGTATTAATGTCAAACTAGTAAAGAGTT  1753

Qy  2157  CGTCAAGTCTCTGGCGCTCACATTCTGCCATTATGCGTGGTCTCGCTTCGCCCAGC  2216
                ||  ||||| || | | |  ||| |  |||  ||   | || |  || | |||
Db  1752  AGAACGTGAGAGGTTTCCATATTTTAGGCGATTAAAGATATGGCAGCATATGGCAGGCTTATTAAAACCTAA  1693

Qy  2217  TGCGGTAACCAAGCTGGTCGTGTCGTGAATTCGATCTGCCAGTGCACGTGCA  2276
                | |||||  ||||| | | | |||| ||| |   ||    | |   |  ||
Db  1692  AGCCGCTTACGAATTGATTGGTGAGTAGTAAAATCGTAGAATTCATCTTCA  1633

Qy  2277  CACCCAGCACTGCGGGGTCGGGTGGTAATGGTTTATTAACATACAAACAATGATGCTGGTGT  2336
                ||  |  |||| | |  |   |  ||| | ||||| | ||||   |  |||||  ||
Db  1632  CACTCATGATACAAGTGGTAATGGTTATTAACATACAAACAATGATGCTTTAATGGTT  1573

Qy  2337  AGATGCTGTTGACGGTGCTTCCGCACCACTGTCTCGGCACCCTCCAGCCATCCCTGTC  2396
                |||  | ||| ||  | |   |||  ||| |||  |   ||| | | |  | ||
Db  1572  CGATATCATTGATACTGCTTCAATGAGTGCTTAACAAGTCAGCCAAGCGCAA  1513

Qy  2397  TGCCATTGTTGCTGCGATTCGCGCACACCCGTCGCGCGATACCGGTTTGAGCCTCGAGGCTGT  2456
                |   ||| |||  || | |   |  |  | |||  | |  |    |||
Db  1512  TTCGTTATATTATGCATTAAAATGGCTTCCACGCGCCACCTTAGAACTGATATTGAAGGTAT  1453

Qy  2457  TTCTGACCTTCGAGCGTACTGGAAGCTGTGCGCGGACTGTACCTGCCATTTGAGTCTCGG  2516
                |||     ||  || || | | | ||| |     |   || |  | ||  ||
Db  1452  GGAGTCACTTAGTCATTATTGGTCAACCGTACGTACTTATTATTCAGACTTTGAAAGTGA  1393

Qy  2517  AACCCCAGGCCCAACCGGTCGCGTCACCCGCGGACAGTTGTCCAA  2576
                |  |  ||  | ||   | |||  ||  | |   |   ||  |||
Db  1392  TATCAAATCACCGAATACTGAAATTTATCAACATGAAATGCCTGGTGGACAGTATTCGAA  1333

Qy  2577  CCTGCGGTGCACAGGCCACCGCACTGGGCCTTGCCTTCGAGCGTCATCGAAGACACA  2636
                | |  || | ||   |  | |  ||| ||  | | | |  || |  |  ||||  ||
Db  1332  TTTAAGTCAACAAGCTAAAAGTTTAGGTTTTAGGCCGTTTGATGAAGATTTGATGAACAAGATAT  1273

Qy  2637  CTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCCATCCTCCCAAGGT  2696
                |||||   |  || ||  |  |     ||| ||  || | | || | | |  || | ||
Db  1272  GTATCGCAGAGTGAATTTCTTATTGGTGATATCGTAAAGTAACACCATCGTCTAAAGT  1213

Qy  2697  TGTTGGGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGCAGACTTTGCTGC  2756
                ||| || | ||  ||| || |  || | |||  ||||  |||    | | || |  || |
Db  1212  AGTTGGTGATATGGCACTTTATATGGTACAAAATGATCTTGATGAACAATCCGTGATTAC  1153

Qy  2757  AGACCCACAAAGTACGACATCCCAGACTTCTGTCATCGCGTCCTGCGGCGGCGAGCTTTG  2816
                || |   ||||| |||   | ||  || ||  | | |   |   ||  | ||   |||
Db  1152  AGATGGCTATAAATTAGATTTCCCAGAATCAGTAGTGTCGTTCTTCAAAGGTGAAATAGG  1093

Qy  2817  TAACCCTCCAGGTGGCTGGCAGCTGCCCGCACCACTGGCACACCGCCTCCGA  2876
                |  |  ||| | |  | |||   |  | || | || |  |   | ||| |
Db  1092  ACAACCTGTAAATGTGTTTAATAAAGATTTACAAGGCCGGTTTATTTTTAAAAGGGCCAAGAGC  1033
```

```
Qy  2877  AGGCAAGGCACCTCTGACGAAGTTCCTGAGGAAGAGCAGGCGCCACCTCGACGCTGATGA  2936
                |||         ||    | |||  |  |||| | ||    | |  ||  |  || | |
Db  1032  ACTAACAGCTCGTCCAGGTGAATATCTAGAGCCCAGTCAGTTGATTTGAAAAGTCCGTGAGTT  973

Qy  2937  TTCCAAGGAACGTCGCAACAGCC--------------TCAACCGCCTGCTGTT  2975
                |||| ||  |  |  ||  |  || |  |||  ||
Db  972   GCTTGAAGAGAGCAACAAGGTCCTGTTACGGAGCAAGAGATATTATTAGTTATGTACTATA  913

Qy  2976  CCCGAAGCCAACCGAGAGTTCCTCGCCACCGTCGGCACCTCGCGCT  3035
                |    |    |  |  | | |  | || ||  || || |   |  ||
Db  912   TCCAAAGTATATGACAATATATCAAACTAGAAATCAATACGGAAACTTATCGTGTACT  853

Qy  3036  GGATGATCGTGAATTCTTCTACGGACTGGTCGAGGGCCGCCGAGACTTTGATCCGCCTGCC  3095
                |     |   |  | |  |     || |  |      || | |   |  |  |  ||
Db  852   TGATACGCCTACATTCTTCTTGGAATGCGTAATG---GTGAAACAGTAGAAATCGAAAT  796

Qy  3096  AGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCGAGCCAGACGATAAGGG  3155
                || |   | || |  | |   |  | |  | ||  | | |    ||| |  |  |||
Db  795   CGATAAAGGTAAACGATTAATTATTAAACTAGAAACGATTAGTGAACCAGATGAAAATGG  736

Qy  3156  TATGCCCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCGAGACGTATTTCACATTAAAGAGTGT  3215
                |   |  ||||  || | | | ||||| || | ||    |  |   || || |  | |
Db  735   TAATAGAACGATTTACTATGCGATGAATGTCAAGCGAGACGTATTTCACATTAAAGAGTGA  676

Qy  3216  CTCCGTTGAGTCTGTCGTTCACCGGACGCAACCGCAGAAAGCCAGATTCCTCCAACAAGGGCCATGT  3275
                |   | | |  | | ||  |   | |  || |   ||   |  | |  || || |  || | |
Db  675   AAATGTGCATACAAATGCGAACGTTAAGCCAAAAGCAGATAAGAGTAATCCAAGTCATAT  616

Qy  3276  TGCTCCACCATTCGCTGGTGTGTCACT--GTGACTGTTGCTGAAGGTGATGAGGTCAA  3332
                | | |   |||| | || | |  ||| ||| | ||| |  ||| || || ||  ||
Db  615   CGGTGCTCAAATGCCAGGTTCAGTAACTGAAGTCAAGGTTAGTGTAGGTGAAACTGTGAA  556

Qy  3333  GGCTCGGAGATGCAGTCGCAATCATCGAGCGCTATGAAGATGAAGCGAACAATCACTGCTTC  3392
                || |   |  | |   |  | ||  |  | ||||  |  || |   || | |  | |  ||
Db  555   AGCTAATCAGCCGTTGCTAATTACTGAAGCTATGAAATGAAAAATGAAACAACAATTCAAGCACC  496

Qy  3393  TGTTGACGGCAAGATTGAACGACGCGTTGTGGTTCCTGCTCAACGAAGGTGGAAGGTTGGCGA  3452
                |||   ||| |  | |  | ||||  || | |   | |    |   |     |  ||| | |
Db  495   ATTTGACGGTGATTAAACAAGTAACTGTAAATATAATGGTGACACAATAGCGACAGGCGA  436

Qy  3453  CTTGATCGTCGTCGTT  3468
                | || |||   | |
Db  435   TTTATTAATCGAAATT  420
```

US-09-070-927A-89

Query Match   14.1%;   Score 488.2; DB 10;   Length 20072;
Best Local Similarity   50.3%;   Pred. No. 2.5e-140;
Matches 1398; Conservative   0; Mismatches 1358; Indels   25; Gaps   7;

```
Qy      86 TCAAAAAGATCTTGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCAC  145
           ||||||||  | || |||| ||| |||| || ||||||| || | |||||||||||
Db    5936 TGAAAAAGGTATTAGTTGCCAATCGTGGCGAAATTGCGATTCGAATTTTAGAGCTTGTA 5877

Qy     146 TCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCACC  205
           || ||||| || ||||||| || ||| |||| |||||||| | |||| | || | |||
Db    5876 CAGAATTAGATATCCGTACAGTAGCGGATTTATGCTGCAGAAGATGAGTATTCTGTTCATC 5817

Qy     206 GCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTCGAAGGCTCACCAGTCAAGGCGTACC  265
           |||   | || ||||||| || ||| ||| |||  ||| |||| |  ||||| |||
Db    5816 GTTTCAAAGCAGATGAAGCGTATTTAGTTGGTAAGGGGAAAAAACCAATTGAAGCCTATT 5757

Qy     266 TGGACATCGATGAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCCGG  325
           ||||||| || || ||||||  | |||| |  |  ||| ||||| || || |||||||
Db    5756 TAGACATTGAAAATTATCCAAATTGCCAAAAATCGGGAGCGATGCCATTCATCCTG 5697

Qy     326 GATATGGCTTCCTGTCTGAAAATGCCCAGCTTACCGGTGAAAACGGCATTA  385
           |||| ||| |||| | ||| |||||| |||||| ||||| | ||| |||||
Db    5696 GTTATGGCTTTTTATCAGAGAACTTAC-GTTCGCCGCCGAACGTTGTGTGAGAGAGGAATCA 5638

Qy     386 CTTTTATTGGCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGTAA  445
           |||| | ||| |||  ||| |   ||| ||||| |||| ||| |||  ||| |||
Db    5637 TTTTTGTCGGACCTAAAACACATCATTTAGATATATTTTTGGCGATAAAATTAAAGCGAAAG 5578

Qy     446 CCGCCCGGAAGAAGGCTGGTCTGCGCCAGTTTGGCCGAATCCACC---CCGAGCAAAAACA  502
           || ||| || |||| || ||  ||||| ||||||  ||  ||||     ||| ||||| |
Db    5577 AAGCAGCTGTAGCTGCGGGGATTGCTTCAGGCTCAGGGGCCAGTAGCACACGG 5518

Qy     503 TCGATGACATCGTTAAAAGCGGTGAAGGCCAGACTTACCCCATCTTTGTAAAGGCAGTTG  562
           || || | ||||| || |||| || ||| |||  ||||| || || |||| ||| |
Db    5517 TGGAAGAGGTTGTAGCTTTTGGTGAAAACATGGCTTTCTTCACCTGATGATTAAAGCTGCTT 5458

Qy     563 CCGGTGGTGCGGACGACGGTCGCGCGGATGCGCCGTTGCCGTTCGCACGGCCATTCGGA  622
           || || || |||| || || ||||  ||| ||| || || || || ||||| |||||
Db    5457 TAGGTGGCGCGGGTGCGCGGTCGCGCGGATGCGCCGTTTCCACGGGCAGCAAGCACGAGGTT 5398

Qy     623 CAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCCGACGGTTCGGTATATGTCGAAC  682
           |||| |||| | |  |||  ||||| |||| ||  || ||||  ||| || | | ||||
Db    5397 ACGAAAGAGCAAAAAGTGAAGGCGAAAGCGAAGCTGAAGACACGGTTTATGTTGAAA 5338

Qy     683 GTGCTGATTAACCCCCAGCACATTGAAGTGCGACATCCTGCGGATCGCACTCGGAGAAG  742
           || ||||||  ||||| |||||| || ||  ||||| || || || || ||||  |||
Db    5337 AGTATTTCTAATCCTAAACATATCGAAGTACAAAAATTTTAGGCGATCATCATGGGACACG 5278

Qy     743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCTTCGTCACCCAAAAAGTGTCCAAA  802
           ||||| |||||||| || || || | |||| |||| |  ||  |||||||| |||||
Db    5277 TCTTGCATTTATTTGAACGTGATTGTTCCGTTCAACGGCGCCACCAAAAAGTGTAGAAG 5218

Qy    1868 CGAATGTAAACATTCAGATGCTGCTTCGCGCGCCAACACCGTGGGATACACCCCGTACC  1927
           ||||||||  | ||| |||||| ||  ||||||||| || || | | |  ||| ||||
Db    1927

Qy     803 TTGCGCCAGCACAGCATTCGAGATCGGTGATCGCTGCATTTGTGCGGATGCGAGTAA  862
           ||  || ||| || |  ||||| || || || || || ||||| ||||||||||||
Db    5217 TCGCACCCATGTGTATCCATGGTTATCCAGGCGCGGGGAACCGCATTGTTCGGCTGCTGTGC 5158

Qy     863 AGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA  922
           |||| || ||| ||||  | ||  |||||||  || || ||  |||||||| |  ||  |
Db    5157 AGTTAATGGCACATGTCGGCGTCGCACGTGAATGCGGGGAACGGTTGAGTTTTTAGT---AGAAG 5101

Qy     923 AGGGCAACCACGTTTCATCGCAATGAACCCACGTATCCAGGTTGAGCACCGTGACTG  982
           ||| ||| |  ||| ||| | |||| ||| |  ||| ||||| |||| |||| | ||
Db    5100 GGGATACAGTTTTATTTTATTTGAAGTGAATCCTGTGTTCAGGTAGAACATACTATCACAG 5041

Qy     983 AAGAAGTCACCGAGGTGGACCTGGTGAAGGCCGCAGATGCGCTTGGCGTCGTTGGTCAACCT  1042
           || || ||| |||| || ||  ||| ||||  |  ||| ||| || |  |||  ||| |
Db    5040 AAATGATTACAGATATTGTGATCTCTCAATTACAAATTGCAAGGGCTTGATT 4981

Qy    1043 TGAAGGA------TTGGGTCTGACCCAAGATAAGATCAAGACCCCACGGTGCAGCACTGC  1096
           ||| ||        ||||||  ||| ||||| |  |||| ||  || ||| ||||| |||
Db    4980 TGCATAAGATATGCATTTGCCAAGAACAAAAACAAATTGACATTAAAAGGCGCGGCTATTC 4921

Qy    1097 AGTGCCGCATCACCACGGAAGATCCAAACACACGGCTTCCCGCCAGATACCGGAACTATCA  1156
           | || ||  || | || |||| ||  | | |||||| || || | |||| || ||  | |
Db    4920 AATGTCGGATTACGACTGAAGATCCCCTGACCAGTTTATGCCAGACACAGGGAAAAATCG 4861

Qy    1157 CCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCT---TGACGGTGCAGCTCAGCTCG  1213
           || | |||| || | ||  | ||| || ||  ||| |    ||| || ||| | ||| |
Db    4860 ATACGTATCGTTCACCAGGTGGTTTTGGTGTTTAGATGTAGAGAATGCTTATTCTG 4801

Qy    1214 GTGGCGAAATCACCCGCACACTTGACTCCATGCTGGTGTCGTGAAAATGACCTGCCGTGGTTCCG  1273
           ||| || || | | || || | |||| | ||  || || || || ||   |||| ||||  |
Db    4800 GCTACGCTGTGACTCCTACTTTGATTCTTTTATTGGTTTAACGTCTGTAACGTCATGGTTTCT 4741

Qy    1274 ACTTTGAAACTGCTGTTGCTCGTGCACAGCGGCGTTGGCTGAGTCACCGTGTCTGGTG  1333
           ||||| || |||| || |||| | || || |  |||| ||  | |||| |||  |||
Db    4740 CTTTTGACACAGCAATCAGTAAAATGCAACCGTGCTTAAAAGAATTTCGGATTCGTGGCG 4681

Qy    1334 TTGCAACCAACATTGGTTCTTGTCGTGCGTTGCGTGCTCGTGCGGGGAGAGAGACTTCACTTCCAAGC  1393
           |||| ||||| ||| || |||  | | ||   ||| || ||||| || ||| || |  ||||
Db    4680 TGAAAACCAAATATTCCGTTTTTACAAAATGTTGTAGCTATCCACGGTTTCAATCTGGGCG 4621

Qy    1394 GCATCGCCACCGGATTTATCGGCGAGTTATGCGGAAACGGCATTA  1453
           ||||| || || || ||  ||| | || ||| | ||| | || |
Db    4620 AAGCCAAAACAACCTTTATTGATAATACACCTGAATTATTTGAATTTCCTCGTATGCGCCG 4561

Qy    1454 ATGAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCAGCCTCAAGCCTCATGGTG  1513
           || || ||| || ||| | ||| | | ||| || || || || || |||| |||| |
Db    4560 ATCGTGGCAATAAAACATGGAAGATAAACATTGGAGAAGAAGTAACGGTCAATGGGTCCCTGGCA 4501

Qy    1514 TGCGTCCAAAGG------ATGTTGCAGCACCAACATCGATAAGCTGCCCAACATCAACGGATC  1567
           ||| |||||||      |||||||| | ||  | |  ||  ||| || |||| |||| ||
Db    4500 TTGAAAGACGGAGAAAAATATTTTGAAGCCCGGCGCTGCCGACAGATATTGAAGTTC 4441

Qy    1568 TGCCACTGCCACGCGGTTCCCGTGACCGCCTTGGCCCGACGGCAGCTTGGAGCGCTTGCTC  1627
           ||||||| || || || ||| |||| ||| ||| |  ||||| |||| ||| || |||
Db    4440 CAGAAAAAGTGATTACAGCTAAAAATATTCTAGACGCTAAAGTCAACCGCCGTCATCG 4381

Qy    1628 GTGATCTCCGTGAGCAGGACGACCTGCAGTTACTGGATCGACACC  1687
           || |||| |||| |||| | | || ||| |  |  ||| ||||||
Db    4380 ATTGGGTTAAGAATCAAGAAAAGTGTTAATGACAGATAACACCACGTTCCGTGATGCGCACC 4321

Qy    1688 AGTCTTTGCTTGCCACCCGAGTCCATTCGCGGACCTCAAGCCTGCGGGCGGCAGAGGCCGTCG  1747
           |||  ||| ||  ||| |||| ||| |   |||| | ||  || |  |||||| |||||| ||
Db    4320 AAAGTTTATTAGCTACTCGTGTGAGAACACAAGATTTTAAAGCAATTGCTGTCTAACTG 4261

Qy    1748 CAAAGCTGACTCCTGAGCTTTTGTCCGTGGCGATCCGACTCGGGGAAC  1807
           ||||| | || || || |  |||||| || || || ||  |||| |||
Db    4260 ATGCAGCCCTGCCTGAGCTGTTTCTAGTGAAAGTGGCGTGGCGCTCACGTTCGATGTTG 4201

Qy    1808 CGATGCGTTTCCTCTTGTGAGGATCCGTGGGACCAGGCTGCGGACAGGCTGCGCGAGGCGATGC  1867
           || |||||| ||  || || |||| |||||| |||| | ||  |||| ||| | |||||||||
Db    4200 CCTATCCGTTCTTAACCGAAGACCCATGGCCACCAAAAGGCCACACGCCGCCCACAAAAAAGTGTAGAAG 4141
```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4413
LENGTH: 3441
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-09-815-242-4413

Query Match   14.0%;   Score 486.8; DB 10;   Length 3441;
Best Local Similarity 48.4%;   Pred. No. 2.5e-140;
Matches 1664;   Conservative   0;   Mismatches 1712;   Indels   60;   Gaps   9;

Qy   84 ATTCAAAAAGATCTTGGTACGAAACCGGCGAATCGCGGTCCGTTCCGTGCAGC 143
            |||||||| |||| ||| |||| ||||||||| |||| ||| |||||||
Db   6 AATAAAAAGTTACTTGTTGCTAACCGTGGAGAAATTGCAATTCGTATATTCAGAGCGGC 65

Qy   144 ACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGGAAGATCGGGGATCATTCCA 203
            ||||||||||||||||||||||||||||||||| | ||| ||||||||||||||||||
Db   66 AGCAGAATTAGACATCAGCACAGTTGCAATTTATTCGAATGAAGACAAAAGTTCATTACA 125

Qy   204 CCGCTCCTTTTGCTTCTGAAGTGTCCGCATTGTACTGATGACTCGAAGGCGTA 263
            ||| |||||| |||||||||||| |||||| |||||| ||| ||||||||||
Db   126 TAGATATAAAGCAGATGAATCCTATTTAGTTGGAAGTGATGATTGATTCCTGCTATTTACCC 185

Qy   264 CCTGGACATCGATGAAATATTACGGTGCAGCTAAAAAATTAAAGCAGAGATGCTATTTACCC 323
            |||||| ||||| ||||||||||| | |||| |||| |||| ||||||||||||
Db   186 TTTAAATATTGAGCGTATCATTGATGTAGCAAAACAAGCAATTGCTGATGCGATTCATCC 245

Qy   324 GGGATATGGCTTCCTGTCTGATAGATCCAGCTTGCTGCGGGAAAACGGCAT 383
            ||||| ||||| ||| |||||||| ||||||| || ||||| ||||||
Db   246 TGGCTATGGATTTTTAAGTGAAAATGAACAATTGCGCGTCGTCACCGGTGATAAGTCGATTATTACCC 305

Qy   384 TACTTTTATTGGCCCAACCCCAGAGGTTCTGCCCAGTTTGGCGGAATCCACCCGAGCAAAAACAT 443
            ||||| ||| ||||| |||||||||||| ||| ||| |||||||||||| ||||||||
Db   306 TAAATTATTGGTCCTCATTTAGAACATTTAGATATGTTTGGAGATAATGTTTAAAGGTAAAGTG 365

Qy   444 AACCGGCGAAGAAGGACGTGGTCTGCCCAGTTTGGCGGAATCGGTCGTGCTGCGGT 503
            |||||||||||||||| ||| ||| |||| |||||||||||||| ||
Db   366 TACAACGGCTATCAAGGCAGATTTACCAGTTATTCCTGGTACAGAGCCGGTCCAATTAAATC 425

Qy   504 CGATGACATCGTTAAAAGCCTGAAGCCCGGAC--TTACCCCATCTTTGTAAAGGCAGT 560
            ||| ||| || |||| |||| ||
Db   426 ATATGAATTAGCAAAAGAATTTGCAGAAGAGAGCTGGTTTCCGCTAATGATTAAAGCCAC 485

Qy   561 TGCCGGTGGTGCGGACGCGGTATGCGTTTGTTGTTCTTCACCTGAGTGAGCTCCGCAAATT 620
            |||| || |||||| | |||| | ||| ||| || | ||| ||

us-09-974-973-1.rnpb

This page contains a nucleotide sequence alignment between two sequences labeled "Qy" (query) and "Db" (database), displayed in two columns of paired rows with position numbers. The alignment is rotated 90° on the page.

Left column (positions ~486–1676):

```
Db   486  AAGTGGTGGCGCGGTAAAGTATGAGAATCGTTCGTGAAGAAAGTGAATTAGAAGATGC  545
Qy   621  GGCAACAGAGACATCTCGTGAAGCTGAAGGGCATTCGGCGACCGTTCGGTATATGTCGA  680
Db   546  TTTCCATAGAGACAAAATCAGAAGCAGAAAAATCATTTGGTAATAGTGAAGTTTACATAGA  605
Qy   681  ACGTGCTGTGATTAACCCCAGCCACATTGAAGTGCGAATCCTTGGCGATCGCACTGGAGA  740
Db   606  AAGATACATTGATAATCCAAAGACATATTGAAGTACAAGTCATAGGTGACGAACATGGAAA  665
Qy   741  AGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCCAAAAAGTTGTCGA  800
Db   666  TATCGTACACTTATTTGAACGTGATTGTTCAGTACAACGTCGTCATCAAAAAGTTGTAGA  725
Qy   801  AATTGCGGCCACAGCAGCCATTTGGATCCAGAACTGCCTGATCGCATTTGTGCGGATGCAGT  860
Db   726  AGTTGCACCATCAGTTGGATTATCACCAACATTACGTCAACGTATTTGTGATGCTGCAAT  785
Qy   861  AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGAATTCTTGGTCGATGA  920
Db   786  TCAATTGATGGAAAATATTAATATCTGATAATGTTATTCATAAATTTCTAGTATCTG-  844
Qy   921  AAAGGGCAACCACGTTTTCATCGAAATGAACCACGTATCCAGGTTGAGCACACCGTGAC  980
Db   845  --GTGACGAATTCTTCTTTATAGAAGTTAACCCTCGTGTACAAGTAGAGACAATTAC  902
Qy   981  TGAAGAAGTCACCGAGGTGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGTGCAAC  1040
Db   903  AGAGATGGTAACAGGAATTGATATTGTTAAGACACAAATTTAGTTGCACGCAGTGCCGA  962
Qy  1041  CTTGAAGGAATTGG--------GTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGC  1091
Db   963  TTTATTTGGTGAAGAAATTAATAATGCCGCAACAAAAAGATATTACACAATTGGGCTATGC  1022
Qy  1092  ACTGCAGTGCGCCATCACCACCGGAAGATCCAAAACAACGGCTTCCGCCCAGATACCGGAAC  1151
Db  1023  CATCCAATGTCGTATTACGTCGTAAATGATTCATGCCGATACTGGAAC  1082
Qy  1152  TATCACCGCGTACCGCTCACCAGGCCGAGCTGGCTTTGTCGTTGACGGTGCCAGCTCAGCT  1211
Db  1083  AATCATTGCTTATCGTTCAAGCGGTGCTGTTGGTGTGTACGTCTTGATGCTGGAGATGGTTT  1142
Qy  1212  C---GGTTGGCGAAATCACCGCACACTTTGACTGCTCCATGCGTGCTGGTGAAAATGACCTGCCGTGG  1268
Db  1143  CCAAGGTGCTGCGAGATATCACCTTATTATGATTCATTACTCGTAAAATATCTTACACATGC  1202
Qy  1269  TTTCCGACTTTGAAAACTGCTGTGTTGCTGCTGCAGCCGCGCCGGTTCATTACGACGGTACGTTC  1328
Db  1203  GATATCATTTAAACAAGCAGAAGAGAAAAAATGTACGTTCATTCATGGCGTATTCG  1262
Qy  1329  TGGTGTTGCAACCAACAACATTGGTTCTTGCGTGCGTTGCTGCGGGGAAGAGGACTTCACTTC  1388
Db  1263  TGGTGTTAAAACTAATAATCCATTCTTAATTAATGTAATAGAAGAATAAAAAGTTCACAAG  1322
Qy  1389  CAAGCGCGATCGCCACCGGATTTATCGGCATCACCGATTTATCGCTCCTTCCCTGGCCTCCACCTGC  1448
Db  1323  TGGTGATTACACAACTAAATTTATTGAAGAAACAACCAGAAACTTTTGCACATTCAGCCGTC  1382
Qy  1449  GGATGATGAGCAGGGACGACGCCATCCTGGATCATTACTTGGCAGATGTCACCCTGAACAAGCCTCA  1508
Db  1383  TCTAGATAGAGAGGTACTAAAACATTAGAATATTAGGTAATGTAACAATTAATGGTTTCCC  1442
Qy  1509  TGGTGT--------GCGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAA  1556
Db  1443  AAATGTTGAGAAACGTCCCAAACCAGACTATGAATTAGCATCAATTCCAACTGTATCTTC  1502
Qy  1557  CATCCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGTGAAGCAGCGTTGGCCCAGC  1616
Db  1503  AAGTAAAATCGCTTCATTT---AGTGGTACGAAACAATTGCTTGATGAAGTAGGTCCAAA  1559
Qy  1617  CGCGTTTGCTCGTGATCTCCGTGACCAGGACGCCACTGGCCCAGTTACTGTGATACCCACTTCCG  1676
Db  1560  AGGTGTAGCTGAACTGGGTTAAAAAGCACGATGATGTCTTACTAACAGATACAACCTTTAG  1619
```

Right column (positions ~1620–2756):

```
Qy  1677  CGATGCACACCAGTCTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGC  1736
Db  1620  AGATGCACCAACCAATCATTATTAGCTACGAGCAGTTAGAACTAAGGATAGATTAATATCGC  1679
Qy  1737  AGAGGCCGTCGCGAAAGCTGACTCCTGAGCTTTGTCCGTGAGGCCTCGGGGCGGGCCGCGAC  1796
Db  1680  ATCCAAAAACAGCGGACGTATTTAAAGATGGTTTCTCACTAGAAATGTGGGGCGCGGTCTAC  1739
Qy  1797  CTACGATGTGGCGATGCGTTCCTCTTTGAGGATCCGTGGGACAGGCTCACGAGCTCGCG  1856
Db  1740  ATTTGATGTGGCATATAATTTCTTGAAGGAAAAACCATGGGAACGACTTGAACGTCTCACG  1799
Qy  1857  CGAGGCGATGCCGAATGTAAACATTCAGATCGCTGCTTCGCGCGCCGCAACACCGTGGGATA  1916
Db  1800  TAAAGCTATTCCAAATGTATTATTCCAAAATGTTGTTACGTGCTTCAAACGACAGTTGGTTA  1859
Qy  1917  CACCCCGTACCCAGACTCCGTGCCGGCCGGTTTGTTAAGGAAGCTGCCAGCTCCGGCGT  1976
Db  1860  TAAAAACTATCCTGATAATGTTATTCATAAATTCGTACAAGAAAGTGCTAAAGCAGGCAT  1919
Qy  1977  GGACATCTTCCGCATCTTCGACGCGCTTAACACGTCTCCCAGATGGCTCCAGCAATCGA  2036
Db  1920  AGATGTCTTTAGAATTTCGATTCATTAAACTGGGTAGATCAAATGAAAGTTGCCAATGA  1979
Qy  2037  CGCAGTCCTGGACACCAACACCGCGGTAGCCGAGGTGGCTATTGCTTATTCTGGTGATCT  2096
Db  1980  AGCAGTACAAGAAGCGGGCAAAATCTCAGAAATGGCTAATATTTGTTATACAGGTGACATTTT  2039
Qy  2097  CTCTGATCCAAATGAAAAGTCTACACCCTGATTACTACCTAAAGATGGCAGAGAGAGAT  2156
Db  2040  AAATCCTGAGCGATCAACAACATTTATACTTTTAGTAGTATTATGTCAAACTAGCTAAAGAGTT  2099
Qy  2157  CGTCAAGTCTGGGCGCTCACATTCTGGCCATTAAGGATATGGCGTCGTCTTCGCCCAGC  2216
Db  2100  AGAACGTGAAGGTTCCATATTTCCATTATTTAGCGATTAAAGATATGCGCAGGCTTATTAAAAACCTAA  2159
Qy  2217  TGCGGTAACCAAGCTGGTCACCGGGAATTCGATCTGCCAGTGCACGTGCA  2276
Db  2160  AGCTGCTTACGAATTGATTGGTGAGTTAAAAGCAGTCGTAGAATTTACCAATCCATCTTCA  2219
Qy  2277  CACCCAGACACTCGCGGCGTGGCCAGTTCGGCCTACCTTTGCTGCAGCTCAAGCTGCGTGC  2336
Db  2220  CACTCATGATCAAGAGTGGTAATGGTTATTAATCGCAGCAATAGATGCTGGTGT  2279
Qy  2337  AGATGCTGTTGACGGTGACGCTGCATTCGGCGCACCACCCGTTGAGCCTCGAGGCGTGT  2396
Db  2280  CGATATCATTGATACTGCTGCTGTTGCTTCAATGAGTGGTTTAACAAGTCAGCCAAGCGCCAA  2339
Qy  2397  TGCCATTGTTGCTGCCATTCGGCGACTCCGAGGCTGTGT  2456
Db  2340  TTCGTTATATTATGCATTAAATGGCTTCCCACGTCACCTTAGAACTGATATTGAAGGTAT  2399
Qy  2457  TTCTGACCTGCGACTGGGGAAGCTGTGCGCGGACTGTACCTATTATTTGAGTCTCGG  2516
Db  2400  GGAGTCACTTAGTCATTCATTATGGTCAACTGTACTGTATTATTCAGACTTTGAAAGTGA  2459
Qy  2517  AACCCCAGGCCCAACCGGTCGCGTCTACCGCCACGAAATCCAGGCGACAGTTGTCCAA  2576
Db  2460  TATCAAATACCGAATACTGAAATTTATCAACATGAAAATGCCTGGTGGACAGTATTCGAA  2519
Qy  2577  CCTGCGTGCACAGGCCACCGCACCCGACCTTGCTCGGGCCTTCTTCGAGCTCATCGAAGACAA  2636
Db  2520  TTTAAGTCAACAAGCTAAAAGTTTAGGTTTAGGCGGAAAGATTTGATGAACTCAAAGATAT  2579
Qy  2637  CTACGCAGCGCGTTAATGACGCTAATGCATCCAATTCAACTGTATCTTC  2696
Db  2580  GTATCGCAGAGTGAATTCTTATTTGGTGATATCGGCCATCGCCATCGTCTAAAGT  2639
Qy  2697  TGTTGCGACCTCGCACTCCACCTGGTTGGCGTGTAGATCCAGCAGACTTTGCTGC  2756
Db  2640  AGTTGGTGATATGGCACTTTATATGTGCAAAAATGATCTTGATGAACAATCCGTGATTAC  2699
```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8263
; LENGTH: 3222
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3222)
US-09-815-242-8263

Query Match 12.3%; Score 428.8; DB 10; Length 3222;
Best Local Similarity 48.1%; Pred. No. 2.6e-122;
Matches 1543; Conservative 0; Mismatches 1607; Indels 60; Gaps 9;

```
Qy    310 GATGCGTATTTACCCGCGGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGC 369
              ||||| || ||  |||  ||||| ||  ||| || || |||||||||  | |||| || |||
Db      4 GATGCGGATTCATCCTGCCTATGGATTTTTTAAGTGAAAATGAAACAATTTGCGGCGTCGTGT  63

Qy    370 GCGGAAAACGGCATTACTACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCCACCGGTGAT 429
              || ||||| ||| |||| ||| |||| ||| ||   |||| ||| |||| ||| | |||||||
Db     64 GCAGAAGAAGGGATTAAAATTTATTGGTCCTCCATTTAGAACATTTAGATATGTTTGGAGAT 123

Qy    430 AAGTCTCGTGCGGTAACCGCCGGAAGAAGGCTGGTCTGCCAGTTTTTGGCGAATCCACC 489
              || ||| ||| | || || |  ||| || || | || ||| ||||  ||
Db    124 AAAGTTAAAGTCGTACAACGGCGTACAACGCAGATTTACCAGTTATCCTGTACAGAC 183

Qy    490 CCGAGCAAAAACATCGATGACACATCGTTA---AAAGCGCTGAAGCGCCAGACTTACCCCATC 546
              | | ||||| | || || || ||||  |||     | |||| ||   |||| || ||  || 
Db    184 GGTCCAATTAAATCATATGAATTAGCAAAAGAATTTGCAGAAGAAGCTGGTTTCCCGCCTA 243

Qy    547 TTTGTAAAGGCAGTTGCCGGTGGTGGCGAAGAAGGCTGGTATGCGCGTTTGTTCTTCACCTGAT 606
              |||| ||| || || |||  || |||| || |     ||   |  |  ||| || || ||
Db    244 ATGATTAAAGCCACAAGTGGTGGCGGCGGTAAAGGTATGAAGAATCGTTCGTGAAGAAAGT 303

Qy    607 GAGCTCCGCAAATTGCAACAAGAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGT 666
              || ||| |||| | | || || |||||  |||||| | || |    || |  || || 
Db    304 GAATTAGAAGATGCTTTCCATAGAGCAAAAATCAGAAGCAGCAGAAAAATCATTTGGTAATAGT 363

Qy    667 TCGGTATATGTCGAACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGC 726
              | |||| || |||| |   || |||||| ||   ||||| ||||||| || ||| ||| 
Db    364 GAAGTTTACATAGAAGAAATACATTGATAATCAAAGCATATATTGAAGTACAAGTCATAGGT 423

Qy    727 GATCGCACTGGAGAAGTTGTACACCTTTATGACGTGACTGCTCACTGCACGTCGTCAC 786
              || || | |||| ||||| |||| ||||| || ||| | |  |    |||| |||||
Db    424 GACGAACATGGAAAATCGTACACTTATTTGAAGTGGATTGTTCAGTACAACGTCGTCAT 483

Qy    787 CAAAAGTTGTCGAAATTGGCCAGCACAGCATTGGATCCAGAACGCGTGATCGCATT 846
              ||||||| | |  | || ||||||||  |||||| | |   || || ||  || ||
Db    484 CAAAAGTTAGAGTTGCACCATCAGTTGATTATCACCACACATTACGTCAACGTATT 543

Qy    847 TGTGCCGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGGCGCGGGAACCGTGGAA 906
              |||| | |||||  |   ||   || | ||| |   |   | || |  |  ||| ||| ||
Db    544 TGTGATGCTGCAATTCAATTGATGGAAATATTAAAATATTAAATATGTGAAATAATGCAGGT 603

Qy    907 TTCTTGGTCGATGAAGAGGGCAACCACGTTTTCATCGAAGTTTTCATCCAGGTT 966
              || || | || ||| || |||| |||| |||| ||  ||||| || || | ||
Db    604 TTTCTAGTATCTG---GTGACGAATTCTTCTTTTATAGAAGTTAACCCTCGTGTACAAGTA 660

Qy    967 GAGCACACCGTGACTGAAGAAGCTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCCTTG 1026
```

```
Qy   2757 AGACCCACAAAGTACGACATCCCAGACTCTGTCATCGCGGTTCTGCGCGGCGAGCTTGG 2816
              || | ||||| | |||| |||||||| ||  || | |  |||| |    |||||| ||
Db   2700 AGATGCTATAAATTAGATTTCCCAGAAATCAGTAGTGTCGTTCTTCAAAGGTGAAATAGG 2759

Qy   2817 TAACCCTCCAGGTGGCTGGCGAACCACTGCGCACCGGTTATTTAATAAAGATTACAAGGCCGGTCCGA 2876
              ||  || | |   |    ||  ||| |    || ||||||||||||||||||| |
Db   2760 ACAACCTGTAAATGTTTTAATAAAGATTTTAATAAAGATGTGCGTTCTTCAAAGGCCAAGAAGC 2819

Qy   2877 AGGCAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGA 2936
              |   |||  |  | ||| || ||| | |||  | |     || |  ||||| |  || 
Db   2820 ACTAACAGCTCGTCCAGGTGGTGAATATCTAGAGCCGGTTGATTTGAAAAAGTCCGTGAGTT 2879

Qy   2937 TTCCAAGGAACGGTCGCAACAGCC-------TCAACCGCCTGCCTGTT 2975
              | || || || |  || || || |        |||  |  ||| |
Db   2880 GCTTGAAGAAGAGCAACAAGTCCGTTACGGAGCAAGAGATATTATTAGTTATGTTATTATA 2939

Qy   2976 CCCGAAGGCCAACCGAAGAGTTCCTCGGCGCACCGTGCCGCTTCGGCCAACACCTCTGCGCT 3035
              ||| |||| |||| |||||| | || || | || ||| |  || || |||| || | | ||
Db   2940 TCCAAAAGTATATGAACAATATTCAAACTAGAAATCAATACGGAAAACTTATCGTACT 2999

Qy   3036 GGATGATCGTGAATTCTTCTACGGACTGGTCGAGGGCGCCGAGACTTTGATCCGCCTGCC 3095
              ||| || | |||||| ||||  || | || | ||| ||  || | |||||| ||| |||
Db   3000 TGATACGGCCTACACTTCTTCTTTGGAATGCGTAAATG--GTGAAACAGTAGAAATCGAAAT 3056

Qy   3096 AGATGTGCGCACCCCACTGCTTGTTCGCGCTGTGTCTGGTCGTCTGGATGCGGATCCTCTGAGCCGACGA 3155
              | || |||| | | ||| ||| ||| || || || || || ||  || ||| || ||  ||| ||| |
Db   3057 CGATAAAGGTAAACGATTAATTATTTAAACTAGAAACGATTAGTGAACCAGATGAAAATGG 3116

Qy   3156 TATGCGCAATGTCGTGTGGCCAACGTCAACGCTCCGCCCAATGCGTGTGCGTGACCG 3215
              |||  || | | | | |||  | ||| | || || ||| |||   ||| | ||  ||
Db   3117 TAATAGAACGATTTACTTACTGCGATGAATGGTCAGCCGAGACCGTATTTATTAAAGATGA 3176

Qy   3216 CTCCGTTGAGTCTGTCAACCGCAACCCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGT 3275
              || ||| ||  |   | || ||||  |||||| || ||||  ||| || ||| |  ||||
Db   3177 AAATGTGCATACAAATGCGAACGTTAAGCGCAAAGCAGATAAGAGTAATCCAAGTCATAT 3236

Qy   3276 TGCTGCACCATTCGCTGGTGGTGTCACT---GTGACTGTTGCTGAAGGTGATGAGGTCAA 3332
              ||| || |  | || |||  | || || |   || || || || |||| |  || || ||
Db   3237 CGGTGCGCAAATGCCAGGTTCAGTAACTGAAGGTAACGGTTAGTGTGTAGGTGAAACTGTGAA 3296

Qy   3333 GGCTGGAGATGCAGTCGCAATCATCCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTC 3392
              |||  ||| | ||| | || | ||    || ||||||| |||| |  ||| |||| | |||
Db   3297 AGCTAATCAGCCGTTGCTAATTACTGAAGCTATGAAGCTATGAAAATGAAAACAACATTCAAGCACC 3356

Qy   3393 TGTTGACGGCAAGATTGAACCGTTGTGGTTCCTGCTGTTGTGGTTGAAGGTGGCGA 3452
              ||||||| | | ||| ||| |  || | ||  |  | |    ||| || || | ||
Db   3357 ATTTGACGGTGTGATTAAACAAGTAACGTAACTGTAAATAATGGTGATCAATAGGCCAGGCGA 3416

Qy   3453 CTTGATCGTCGTCGTT 3468
              ||| || | | |  ||
Db   3417 TTTATTAATCGAAATT 3432
```

RESULT 11
US-09-815-242-8263
; Sequence 8263, Application US/09815242
; Patent No. US2002061566A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21

```
Db   661  GAGCATACAATTACAGAGAGTGGTAACAGGAATTGATATTGTTAAGACACAAATTTTAGTT  720
Qy  1027  GCTGCTGGTGCAACCTTGAAGGAATTGG-------GTCTGACCCAAGATAAGATCAAG    1077
Db   721  GCAGCAGGTGCCGATTTATTTGGTGAAGAAATTAATATGCCGACAACAAAAGATATTACA  780
Qy  1078  ACCCACGGTGCAGCACTGCAGTCGCGCATCCAGTGCCATCCACCGGAAGATCCAAACACGGCTTCCGC  1137
Db   781  ACAATGGGAACAATCATTGCTTATCGTTCAAGCGGTGGCTTGGTGTACCGTCTTGAT  840
Qy  1138  CCAGATACCGGAACTATCACCGGCTACCGGCTCACCAGGCGGAGCTGGCCTTCGTTCGTTGAC  1197
Db   841  CCGGATACTGGAACAATCATTGCTTATCGTTCAAGCGGTGGCTTGGTGTACCGTCTTGAT  900
Qy  1198  GGTGCAGCTCAGCTC---GGTGGGCGAAATCACCCGCCACACTTTGACTCCATGCTCGGTGAAA  1254
Db   901  GCTGGAGATGGTTTCCAAGGTGCTGAAACTCGTTCGTCGTGCACAGCGGCGCGTTGGCT  960
Qy  1255  ATGACCTGCCGTGGTTCCGACTTTGAAACGTCGTGTGCTCGTGCACAGCGGCGCGTTGGCT  1314
Db   961  TTATCTTACACATGCGGATATCATTCATTTAAACAAGCAGAGAAGAAAAAATGGTACGTTCATTACGA  1020
Qy  1315  GAGTTCACCCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCCGTTGCTGCGCGGAA  1374
Db  1021  GAAATGCGTATTCGTGGTGTTAAAACTAATAATCCCATTCTTAATTAATGTAATGAAGAAT  1080
Qy  1375  GAGGACTTCACTTCCAAGCGCATCGACGATTTATCCGCGATTATCGGCGATCACCCACCTCCTT  1434
Db  1081  AAAAAGTTCACAAGTTGGTGATTACAACAATTAATTGAAGAAACACCAGAACTTTTC  1140
Qy  1435  CAGGGCTCCACCTGCCGATGAGAGCAGGGACGGCCATCCTGGATTACTTCTGGCAGATGTCACC  1494
Db  1141  GACATTCAGCCGTCTCTAGATAGAGGACTACTAAAACATTAGAATATATAGGTAATGTAACA  1200
Qy  1495  GTGAACAAGCCTCATGGTGT-------GCGTCCAAAGATGTTGCAGCACCAATC  1542
Db  1201  ATTAAATGGTTTCCCAAATGTTGAGAAACGTCCGAAAACGAGCACTATGAATTAGCATCAATT  1260
Qy  1543  GATAAGCTGCCAACATCAAGGATCTGCCTACTGCGCCACGGTTCCCGTGACCGCCTGAAG  1602
Db  1261  CCAACTGTATCTTCAAGTAAAATCGCTTCATT---TAGTTGGTACGAAACAATTGCTTGAT  1317
Qy  1603  CAGCTTGGCCCCAGCCGCGTTTGCTCCTGATCTCCGTGAGCAGGACGCACTGGCAGTTACT  1662
Db  1318  GAAGTAGGTGCAAAAGGTGTAGCTGCCATTTGTGATGTGGCATATAAAGCAGGATGATGTCTTACTAACA  1377
Qy  1663  GATACCACCTTCCGCGATGCACACACCAGTCTTGCTGCCACCCCGAGTCCGCTCATTCCGCA  1722
Db  1378  GATACAACCTTTAGAGATGCGCCACCACCATTAGCTGCTTAGAGAGTTAAGAACTAAGGAT  1437
Qy  1723  CTGAAGCCTCGCGGCAGAGGGCCGTGCCAAAGCTCCTCGAGCTTTTGTCCGTGGAGGGC  1782
Db  1438  ATGATTAATATCGCATCCAAAAACAGCGGAGACGTATTTAAAGATGGTTTCTCACTAGAAATG  1497
Qy  1783  TGGGGGGCGCGCGACCTACGATGTGGCGATGCGTTTCCTCTTTTGAGGATCCGTGGGACAGG  1842
Db  1498  TGGGGGCGGTGTGCTACATTTGATGTGGCATATAATAATTTCGATTCATTAAACTGGGTAGATCAAATG  1557
Qy  1843  CTCGACGAGCTGCGCGAGGCGATGCGAATGCGAATGTAAACATTCAGATGCTCGCTTCGCGCGGCCGC  1902
Db  1558  CTTGAACGTCTCTACGTGAAAGCTAAAGCTAAAATGTTGTTACGTGCTTCA  1617
Qy  1903  AACACCTGGGATACACCCGTACCCGTACCTCCGCGCGCGTTTGTTAAGGGAAGCT  1962
Db  1618  AACGCAGTTGGTTTATAAAACTATCCTGATAAATGTTATTCATAAATTCGTACAAGAAAGT  1677
Qy  1963  GCCAGCTCCGGCGCGACCATCTTCCGACGCGCTTAACGACGTCTCCCAGATGTG  2022
Db  1678  GCTAAAGCAGGCATAGAGATGTGCTTTAGAATTTTCGATTCATTAAACTGGGTAGATCAAATG  1737
Qy  2023  CGTCCAGCAATCGACGCGAGTCCTGGAGAACCAACACCGGTAGCCGAGGTGGCTATGGCT  2082
Db  1738  AAAGTGCCAATGAAGCAGTACAAGAGAAGCGGGCAAAATCCAGAAGGTACTATTTGTTAT  1797


Db  2083  TATTCTGGTGATCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAG  2142
Qy  1798  ACAGGTGACATTTAAAATCCTGAGCGATCAAACATTTATACTTAGAGTATTATGTCAAA  1857
Db  2143  ATGGCAGGAGGAGATCGTCAAGTCTGCGCTCAAGTCTCTGGCCATTAAGGATATGGCTGGT  2202
Qy  1858  CTAGCTAAAGAGTTAGAACGTGATTAAAGATGATTAAAGATATGGCAGGC  1917
Db  2203  CTGCTTCGCGCCAGCTGCGTAACCAAGCTGGTGCATGCGCCGTGAATTCGATCTG  2262
Qy  1918  TTATTAAAACCTAAAGCTGCTTACGAATTGATTGGTGAGTTAAAAGCAGCTGTAGATTTA  1977
Db  2263  CCAGTGCACGTGCACACCCACACTGCGCGGTGGCCAGTTGGCTACCTTGTCTGCA  2322
Qy  1978  CCAATCCATCTTCACACTCATGATCAATGTAATGGTTTTATTAACGTACAAACAAGCA  2037
Db  2323  GCTCAAGCTGGTGCAGATGCGTGTTGACGATGCGTTCCGCCACCACTGTCTGGCACCACCTCC  2382
Qy  2038  ATAGATGCTGGTGTCGATATCATTGATACTGCTGTTTGCTTCAATGAGTGGTTTAACAAGT  2097
Db  2383  CAGCCATCCCTGTCTGCCATTGTGCTGCATTCGCGCACCCGCTGCGATACCGGTTTG  2442
Qy  2098  CAGCCAAGCGCCAATTCGTATTATTATGCATTAAATGCTTCCCAGTCACCTTAGAACT  2157
Db  2443  AGCCTCGAGGCGTGTTTCTGACCTCGAGCGTACTGGAAGCTGTGCGCGGACTGTACCTG  2502
Qy  2158  GATATTGAAGGTATGGAGTCACTTAGTCAACTGTACGTACTTATTATTATTCA  2217
Db  2503  CCATTTGAGTCTGGAACCCCAGGCCAACGGTCGGCCGTCTACCGCCACGAAATCCCAGGC  2562
Qy  2218  GACTTTGAAAGTGATATCAAATCAACATCACCGAATACTGAAATTTATCAACATGAAATGCCTGGT  2277
Db  2563  GGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCTGCGGCCTTCCGAG  2622
Qy  2278  GGACAGTATTCGAATTTAAGTCAACAGCTAAAGTTTAGGTTTAGGCGAAAGATTTGAT  2337
Db  2623  CTCATCGAAGAACAACTACGCAGCCGTTAATGATGAGATGCTGGGACGCCCAACCAAGGTCACC  2682
Qy  2338  GAAGTCAAAGATATGTATCGCAGAGTGAATTTCTTATTTTGGTGATATCGTAAAAGTAACG  2397
Db  2683  CCATCCTCCAAGGTTGTTGGCGCACTCCGCACCCTGGTTGGTGCGGGGTGTAGATCCA  2742
Qy  2398  CCATCGTCTAAAGTAGTTGGTGTGATATCGGCCACTTTATATGGTACAAAATGATCTTGATGAA  2457
Db  2743  GCAGACTTTGCTGCAGACCCACAAAAGTACGACATCCCAGATACGACATCTGTCATCCGGTTCCTG  2802
Qy  2458  CAATCCGTGATTGATTACAGATTACAGATGGCTATAATTTAGAATTAGGATT  2517
Db  2803  CGCGGCGAGCTTGGTAACCCTCCAGGTGCTGCTGCCGCACCCGCGCACTG  2862
Qy  2518  AAAGGTGAAATAGGACAACCTGTAAATGGTTTTAATAAAAGATTACAAGCGGTTATTTTA  2577
Db  2863  GAAGGCCGCCTCCGAAGGCAAGGCACCTCTGAGGATCCGGAAGTTCCTGAGGAAGAGCAGCGGCCAC  2922
Qy  2578  AAAGGCCCAAGAGCACTAACAGCTCGTCGCAGGTGAATTACTGAGACCCGGTTGATTTGAA  2637
Db  2923  CTCGACGCTGATGATTCCAAGGAACGTCGCAACAGCC-------------TC  2961
Qy  2638  AAAGTTCCGTGAGTTGCTTGAAGAGAGCAACAAGGTCCTGTTACGCGGACCGCCAAGCAAGATATTATT  2697
Db  2962  AACCGGCCTGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTTCGGCTTCGGC  3021
Qy  2698  AGTTATGTATTATCCAAAAGTATATGAACAATATATTCAAACTCAATACCGGA  2757
Db  3022  AACACCTCTGCGCTGGATGATCGAATCTCTACGGACTCGTCGAGGGCCCGAGACT  3081
Qy  2758  AACTTATCGTTACTTGATACGGCCTACATTCCTTCTTTTGGAATGCGTAATGGT---GAAACA  2814
Db  3082  TTGATCCGGCCTGCCCAGATGTGCGCCCCACTCCTTGTTCGCCGATGGCGATTCTCAG  3141
Qy  2815  GTAGAAATCGAAATGTGAAATGCGGGCAAAATCTCAGAAATCCAGAGAACGATTAGTTGAA  2874
```

```
Qy   3142  CCAGACGATAAGGGTATGCGCCAATGTTGTGCCAACGTCAACGGCCAGATCCGCCCAATG  3201
           |||||||||||  ||||  ||||  |||  |||||||||  ||||||||||||||||||
Db   2875  CCAGATGAAAATGGTAATAGAACGATTTACTATGCGATGAATGGTCAAGCGAGACGTATT  2934

Qy   3202  CGTGTGCGTGCACCGCTCCGTTGAGTCTGTCACCGGAACCGCAGAAAAGGCAGATTCCTCC  3261
           ||||  |  |||  |||||||||  |||  ||||  |||  ||||||  ||  |||||||
Db   2935  TATATTAAAGATGAAAAATGTCATACAAATGCGAACGTTAAGCAAAAGCAGATAAGAGT  2994

Qy   3262  AACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTGTTGTCACT---GTGACTGTTCGTGAA  3318
           || ||  |  || |||||||||  |||||  ||  |||||  |     |||||  ||||
Db   2995  AATCCAAGTCATATCGGTGCGCAAATGCCAGGTTCAGTAACTGAAGTCAAGGTTAGTGTA  3054

Qy   3319  GGTGATGAGGTCAAGGCTGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCA  3378
           ||||||||  |  ||  |||  |||  |||||||  ||||||  ||  ||||  ||||
Db   3055  GGTGAAACTGGAAAGCTAATCAGCCGTTGCTAATTACTGAAGCTATGAAAATGGAAACA  3114

Qy   3379  ACAATCACTGCTTCTGTTGACGGCAAGATGAACCGTTGGTTCCTGCTGCCAACGAAG  3438
           |||  |  |  |  ||||  |||  |  ||  ||||  ||||  |  |||||||||
Db   3115  ACAATTCAAGCACCATTTGACGGGTGTGATTGAAACAAGTAACTGTAAATAATGGTGATACA  3174

Qy   3439  GTGGAAGGTGGCGACTTGATCGTCTCGTT  3468
           ||   ||||  |  ||||  |  |||||
Db   3175  ATAGCGACACGGCGATTTATTAATCGAAATT  3204
```

```
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-767-479-7

Query Match         8.8%;     Score 305;   DB 10;     Length 1362;
Best Local Similarity  53.5%;   Pred. No. 4.7e-84;
Matches  713;   Conservative  0;  Mismatches 605;   Indels 15;   Gaps  3;

Qy    85  TTCAAAAAGATCTTGGTAGCAAACCGGCGGCAAATCGCGTCCGTTCCGTGCCAGCA  144
          ||||||||  |||||||  ||  |||||||||  |||||||||||||||||||||
Db     7  TTCAACAAGATCCTGATCGGCCAATCGGCGGCAAATCGCCCTGCGCAATCTCCGCACTTGT  66

Qy   145  CTCGAAACCGGTGCAGCACGGTAGTATTACCCCCGTGAAGATCGGGGATCATTCCAC  204
          ||  |||||||||  |||||||  ||  |||  ||||||  |||||||||||||
Db    67  GAAGAACTCGGGATCGGATGGCACGCACGTTCACTCCACTGGATGCAACGCGCTCCAT  126

Qy   205  CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACTGGTACTGAAGCTCACCAGTCAAGGGGTAC  264
          ||  |  ||  ||||||||  ||||||||  ||  |  |||||  ||  |||||||||
Db   127  GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGGCGCAGCAGCAAAGCTAT  183

Qy   265  CTGGACAATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCCG  324
          ||  |||||  ||  |||  |  ||  |  ||  |||  ||  |  ||||  ||  ||||
Db   184  CTCAATATCCCCAACATCATTGCGCGCCCTGACCCGTAATGCAAGCGCCATTCACCCCC  243

Qy   325  GGATATGGCTTCCTGCTGAAAATGCCCAGTTGCCCGAGTGCGCGGGAAAACGGCATT  384
          |||||  |||  |||  ||  |  ||  |||  |||  |  ||||||||  |||||
Db   244  GGCTATGGCCTTCTTGCTGGCGAATGCCCGGCTTTGCAGAAATCTGCCGCCGATCACCATCTC  303

Qy   385  ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTG---G  441
          ||||||||||  ||||||||||  |||  ||||||  |||||||||||  ||||||
Db   304  ACCTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGCGGCGATAAATCACCGGCTAAG  363

Qy   442  GTAACCGCCCGAAGAAGGCGGTCGGCGTTCCGCAGATTCCGGCGAGTGCACCCCGCAAAAAC  501
          ||  |  ||||  ||||  |||  ||  ||||  |  ||  ||||||  |  ||||  ||
Db   364  GAAACAATGCAGCGGGTCGGCGTTCCGGCGGCAGTTCGGGCAGTGCAGTGACGGTCTGCTGACGGAT  423

Qy   502  ATCGATGACATCGTTAAAAGGCGCTGAAGGCCGAGCACTTACCCCATCTTTGTAAAGGCAGTT  561
          ||  ||  ||||||  ||||||  |  |||||  |  ||||  ||  ||||||  ||||
Db   424  GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATCATGAGCCATCGAAAGGCACG  483

Qy   562  GCCGGTGGTGCGGACGCGGTATGCGGTGCGGTATGCGGCTTGTTCTTCACCTGATGAGCTCCGCAAATTG  621
          ||  |  |||||  ||||  |||||  |||  ||  |  |||  ||||  ||||||
Db   484  GCGGGGGCCGCGGTGGTCGCGGTATGCGGTATGCGGTGAGCCTGTGTGCGAGATCTGAAAAACTG  543

Qy   622  GCAACAGAAGCAATCTCGTGAAGCTCGAAGCGGTCGAAGCGGTTCGGCGACGGTTCGGTATATGTCGAA  681
          |||  ||||||||||  ||  |  ||  |||  ||||  ||  ||||  |||  ||
Db   544  TTCCTTGCTGCCAAGGAGAAGCGGAGGCAGCTTTTGGAATCCAGGACTGTATCTCGAA  603

Qy   682  CGTGCTGTGATTAACCCCCAGCACATTGAAGTGCACACTCCTGGCGATCGCGACTGCGAGAA  741
          ||  |  |  |||  |  |||||||  |||||||  |  ||  |  |||||||||||
Db   604  AAATTTTATCGATCGGCCACGTTGAATTTGAATTTCAGATTCGAATTTCAGATTTGGCGCGATCGTTACGGCAAT  663

Qy   742  GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCTGCTGTCACCAAAAGTGTCGAA  801
          ||  |  ||||  |  ||  |  |||||  |||  |||  ||  |||||  ||||
Db   664  GTAGTGCATCTAGGCGAGGCGCGATTGCTCCATTCCATTCAACGTCGTCACCAAAAGCTGCTCGAA  723

Qy   802  ATTGGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTTGTGCCGGATGATGCAGTA  861
          ||  |||  |  |  |  ||||||  ||||  |||||||  |||  |||  ||||  ||
Db   724  GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGCGCAGAAAATGGGCGGATGCGCCGCCGTC  783

Qy   862  AAGTTCGCCGCTCCATTGGTTACCAGGGCGGGGAACCGTGGGAATTCTTGGTCGATGAA  921
          |||  |||||  ||  ||||  ||  ||||||||||  |||||||  ||  ||||  ||||
Db   784  AAAGTCGCTCAAGCGACGATCGGCTACGATCGGTGCCGCGGCACCGTGGGAGTTTCTGGTCGATGCG  843

Qy   922  AAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT  981
          |||||||  |  |||  |  ||  |||  ||  |  ||  |||||  ||||  ||||
Db   844  ACCGGACAACTTCTACTTCATGGAGATGAATAACCCGATCCAAGTCCAAGTCGAGCAATCACCACA  903

Qy   982  GAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGGCTTGCGCTGTCGGTGCAACC  1041
          |||||  ||  ||  ||  ||||  |||  |  |||  ||  |  ||  ||||  |
Db   904  GAAATGATTACGGACGTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG-----  958

Qy  1042  TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGAACCCAGCCGTGCAGCACTGCAGTGC  1101
```

```
Db   959 ----AAGCCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014

Qy    89 AAAGATCTTGGTAGCAAACGCGGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCACTCG  148
         |   || | |||| ||| ||||||||||||| || || || | |||| | ||| |  |
Db     8 ATAAAATTGTTATTGCCAACGCGGCGGCGAGATTGCATTGCGTATTCTTCGTGCCTGTAAAG   67

Qy   149 AAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGATCGGGATCATTCCACCGCT  208
         ||| || || || ||| ||| ||| |||| ||||| || ||| ||| ||| ||||
Db    68 AACTGGGCATCAAGACTGTCGCTGTGCACTCCAGCCGATCGGATCGCGATCTAAAACACGTAT  127

Qy   209 CTTTGCTTCTGAAGCTGTCCCCATTGGTACTGAAGGCTCACCAGTCAAGGCGTACCTGG  268
         || | ||| | |||| | |||| || |||||| |||||   ||
Db   128 TACTGGCAGATGAAACGGTCTGTATTGGCCCTGTCCGT---CAGTAAAAAGTTATCTGA  184

Qy   269 ACATCGATGAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGCTATTACCCGGGAT  328
         ||||| || || ||| ||| ||   | |  ||||| || || | || |  || ||||
Db   185 ACATCCCGGCAATCATCAGCGCGCGCTGAAATCACCGGCGTAGCAATCATCCGGGTT  244

Qy   329 ATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGGCAGTGCGCGAAAAACGCATTACTT  388
         || | ||| | ||||| || || ||| || | || ||| |||| |||| ||| || |
Db   245 ACGGCTTCCTCTCCGAGAACGGCCAACTTGCCGAGCAGGTTGAACGCTCCGGCTTTATCT  304

Qy   389 TTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCGTGCGGTAACCG  448
         ||   ||| ||| | || ||| | |||| || ||||| ||||| | || |  ||||
Db   305 TCATTGGCCCGAAAGCAGAAACCATTCGCCTGATGGGCGACAAAGTATCCGCAATCGCGCG  364

Qy   449 CGCGGAAGAAGGCTGGTCTGCGCCAGTTTGGCGGAATC------CACCCCGAGCAAAAACA  502
          | | ||  |||| | || || |||| | | | ||||      ||||| ||| | | ||
Db   365 CGATGAAAAAAGCGGGGCGTCCCTTACCGGGGTTCTGACGGCCCGCTGGCGGCGACGATA  424

Qy   503 TCGATGACATCGTTAAAAGCCTGAAAGCCAGACTTACCCCATCTTGTAAGGCAGTTG  562
          || || || | |  |||| ||| | |  |||| | ||  ||||||  ||| |||
Db   425 TGGATAAAAACCTGCCATTGTAAACGCATTGGTTATCCGTGATTATCAAAGCCTCCG  484

Qy   563 CCGGTGGTGGCGGACGCGGTATGCGCCTTGTTTTTCTTCACCTGATGAGCTCCGCAAATTGG  622
         ||| |||||| ||||||| || ||||| || | ||||  |||| | || || |||| |||
Db   485 GCGGCGGCGGCGGTCGCGGTATGCGGTAGTGCCGGCGACGCTGGCACAATCCA  544

Qy   623 CAACAGAAGCATCTCGTGAAGCTGAAGCGGGCATTCGGCGACGGTTCGGTATATGTCGAAC  682
          || | || || |  ||||| || ||| |||  | || || |||  | || ||||| |||
Db   545 TCTCCATGACCCGTGCGGAAGCGAAAGCTGCTTTCAGCAACGATATGGTTTACATGGAGA  604

Qy   683 GTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATCTTGGCGATCGCACTGCGAGAAG  742
         ||| || | |  || || ||||||| || |||||| ||| | || || ||| || |||
Db   605 AATACCTGGAAAATCCTCGCCACCGTGAGATTCAGGTACTTGCTGCTCGACGGGCCAACG  664

Qy   743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCGCACCAAAAAGTTGTCGAAA  802
         || | || |||   || || ||||| || |||| |||  || |||  |||| ||| |||
Db   665 CTATCTATCTGCGGGAACGTGACTGCTTCCATGCAACGCCGCACCCAGAAAGTGGTCGAAG  724

Qy   803 TTGCGGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCCGCATTTGTGCGGATGCAGTAA  862
         || |||| || || ||||| | || ||| || |||| | |||  || ||| || |||||
Db   725 AAGCGCCAGCACCGGCATTACCCGGACTCGTCGTACATCGGCGAACGTTGCGCTA  784

Qy   863 AGTTCTGCCGCTCCATTGGTTACCAGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAA  922
         ||||| || || |  ||  ||| ||| || | |||| | ||| |||  |||| | |||
Db   785 AAGCGGTGTTGTTGATATCGGCTATCGGCGGTGCAGTGCTCGAGTTCCTGTTC---GAAA  841

Qy   923 AGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCGGTGAGCACCGTGACTG  982
         || || || ||| ||| || |||| || ||| ||| ||| |||| || |||| ||
Db   842 ACGGCGAGTTCTATTTCATCGAAATGAACAACACCCGGTATTCAGTAGAACACCCGGTTACAG  901

Qy   983 AAGAAGTCACCGAGGTGACCTGGTGAAGCGGCAGATGCGCTTGGCTGCTGTGTCGTCAACCT 1042
         || ||||| ||||| || |||| | ||| | ||| || | || |  ||||| || |||||
Db   902 AAATGATCACCGGCGTTGACCTGACCGTATCAAAGAGAACAGCGCGTATCGCGCCGTCAACCGC  961

Qy  1043 TGAAGGAATTGGGTCTGACCCCAAGATAAGATCAAGACCACGGTGCAGCACTGCAGTGCC 1102
         |||| | |  |   |||| | | |||  ||||| |  || |||| |||| || ||| |
Db   962 TGTCG-------ATCAAGCAAGAAGAAGTTCACGTTCGCGCCGCCATGCGGTGGAATGTC 1012

Qy  1103 GCATCACCACGGAAGATCCAAACAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGT 1162
         |  || || |  ||| | ||| ||||  || ||    |||| ||   |||  | ||||| |
Db  1013 GTATCAACGCCGAAGATCCGGAACA---CCTTCCTGCCAAGTCCGGGCAAAATCACCCGTT 1069
```

```
Db   959 ----AAGCCTGCGCTTCCGGCAAGCCGATATTCAACTGCGCGGCCATGCGATCGAATGC 1014
Qy  1102 CGCATCACCACGGAAGATCCAAACAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1161
Db  1015 CGTATCAATCGGGAAGATCCGGAATACAATTCCGGCCCGAATTCCGGCCGCATTACAGC 1074
Qy  1162 TACCGCCTCACCAGGGGAGCTGGCGTTCTTGCAGGTGCAGCTCAGCTCGGTGGCGGAA 1221
Db  1075 TATTTACCGCCGGCCGGCCCGGGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134
Qy  1222 ATCACCGACACACTTTGACTCCATGCTGGTTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1281
Db  1135 ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTGTCGGGGTGCAACACGGGAA 1194
Qy  1282 ACTGCGTGTTGCTCGTGCACAGCGCGCGCGTTGGCCTGGCTGTCGAGTTCACCGTGTTGCAACC 1341
Db  1195 GAGGCGATCGCGCCGATGCACGCGGTGCTCTCGCGGGAATGCGCCATCACCGGCTTGCCCGACG 1254
Qy  1342 AACATTGGTTCTTGCGTGCGTTGCCTGGCGGAGGAGGACTTCACTTCCAAGCCCGCATCGCC 1401
Db  1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGGAACTCTAT 1314
Qy  1402 ACCGGATTTATCG 1414
Db  1315 ACCAACTTTGTTG 1327
```

RESULT 13
US-09-815-242-6233
; Sequence 6233, Application US/09815242
; Patent No. US2002006156A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6233
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1350)
US-09-815-242-6233

Query Match   8.7%; Score 302.2; DB 10; Length 1350;
Best Local Similarity   54.9%; Pred. No. 3.4e-83;
Matches 714; Conservative 0; Mismatches 563; Indels 24; Gaps 5;

```
Qy   1163 ACCGGCTCACCAGGCCGGAGCTGGCGTTCGTCTTGACGGTCAGCTCGGTGCGTGCGGCGAAA 1222
          ||||| |||||| |||||  |||    |||||| ||||||||||||||||  ||||||||||
Db   1070 TCCACCGACCTGGCCGGTTTTGCGTACGTTGGGAGTTCATATCTACGCGGCTCACCG 1129

Qy   1223 TCACCCGCACACTTTGACTCCATGCTCTGGTGAAAATGACCTGCCGTTGTTCCGACTTTGAAA 1282
          |||      |||||||||| ||  |||| || |||||||||||||||||||||||||||||
Db   1130 TACCGCCGTACTATGACTCAATGATCGGTGTAAGCTGGATTTGCTACGGTGAAAACCGTGACG 1189

Qy   1283 CTGCTGTTGCTGCGCACAGCGCGCGGTGCGTTGGCTGAGTTCACCGTGTTGTTGCAACCA 1342
          |||   |||||| |||| |  |||||| |||    ||||||||||| ||||||||||||||
Db   1190 TGGCGGATTGCCCGCATGAGAGAATGCCCTGCAGGAGCTGATCATCGACGGTATCAAAACCA 1249

Qy   1343 ACATTGGTTTCTTGCGTGCGTTGCTGCGCGGAAGAGGACTTC 1383
          |||||||||||||||||| |||||||||||||||||||||||
Db   1250 ACGTTGATCTCGAGATCCGCCATCATGAATGACGAGAACTTC 1290


RESULT 14
US-09-767-479-5
; Sequence 5, Application US/09767479
; Patent No. US2001003665A1
; GENERAL INFORMATION:
;   APPLICANT: Haselkorn, Robert
;      Gornicki, Piotr
;   TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
;          METHODS OF USE
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: United States of America
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/767,479
;     FILING DATE: 22-Jan-2001
;     CLASSIFICATION: Unknown
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/468,793
;     FILING DATE: <Unknown>
;     APPLICATION NUMBER: US SN 07/956,700
;     FILING DATE: 02-OCT-1992
;     APPLICATION NUMBER: PCT/US93/09340
;     FILING DATE: 30-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Kitchell, Barbara S.
;     REGISTRATION NUMBER: 33,928
;     REFERENCE/DOCKET NUMBER: ARCD:152/KIT
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (512) 418-3000
;     TELEFAX: (713) 789-2679
;     TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 3065 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
;   SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-767-479-5

  Query Match       8.5%;  Score 297; DB 10;  Length 3065;
  Best Local Similarity 54.4%;  Pred. No. 2.2e-81;
  Matches  691;  Conservative   0;  Mismatches  565;  Indels  15;  Gaps    4;
```

```
Qy     85 TTCAAAAAGATCTTCGTAGCAAACCCGGCGAAATCGCGGTCCGTCGTTTCCGTCGCAGCA 144
          ||| ||||| ||  ||  |  ||  |  |||  |||||||||||||||| |  | ||||| |
Db   1314 TTTGACAAAATTATTAATTGCCAATCGGGGAGAAATAGCGTCTGCGCATTCTCCGCGCTGT 1373

Qy    145 CTCGAAACCGGTGCAGCCACGTAGCTATTCACCCCGTGAAGATCGGGGATCATTCCAC 204
          |  |||| ||||||| || ||| |  || | |||  |||||| | |||||||||| ||
Db   1374 GAGGAAATGGGGATTGCGACGATCGCAGTTCATTCGACGTGTTGACCGGAATGCTCTTCAT 1433

Qy    205 CGCTCTTTGCTTCGAAGCTGTCCGCATTGGTACTGTACTGAAGGCTCACCAGTCAAGGGGTAC 264
          ||||  ||||| | |||||| | || |||  |||||||| ||| ||||||| ||||||| |
Db   1434 GTCCAACTTGCTGACGAAGCGGGTTTGTATTGGC---GAACCTGCTAGCGCTAAAAGTTAT 1490

Qy    265 CTGGACATGCGATGAAATTATCGTGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCCG 324
          ||||| |   |  |  ||| |  ||||||  ||| || | |||||  |  ||| || ||||
Db   1491 TTGAATATTCCCAATATTATTGCTGGGGGCTTTAACGGCACGTGCTATTCATCCT 1550

Qy    325 GGATATGGCTTCCTGCTGCTGAAAATGCCCAGCTGCCCGGAGTGCGCGGAAAACGGCATT 384
          ||||||||||   || ||| || || ||  || ||||| |||||||| |||| || || |
Db   1551 GGGTATGGCTTTTTATCTGAAAATGCCGAAAATCTGTGTGACCATCACATT 1610

Qy    385 ACTTTTATTGGCCCACCCCAGAGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTA 444
          |  |||| ||||||| |||||| | | |||| |  || || |||  |  ||| ||| |
Db   1611 GCATTCATTGGCCCCACCCCAGAAGCTATCCGCCTCATGGGGGACAAATCCACTGCCAAG 1670

Qy    445 ACCGCGCGAAGAAGGCTGGTCTGCCAGTTTGGCCG--AATCCACCCGAGCAAAAACA 502
          || || ||| | |||||  | ||||| ||||| ||  ||||| |||| |||| | |||
Db   1671 GAAACCATGCAAAAGCTGGTGTACCGGTGTAGTGAAGGTTTGGTAGACA 1730

Qy    503 TCGATGACATCGTTAAAAGGCTGAAGGCA-GACTTACCCCATCTTGTAAAGGCAGTT 561
          || || | |||   ||  ||| || ||| | |||  ||| |||| ||| ||||| ||
Db   1731 GAGCAAGAAGGATTAGAACTGGCAGAAGATAATGGCTACCCAGTGATGATCAAAGCCACG 1790

Qy    562 GCCGGTGGTGGCGGACGCGGTATGCCGGTTATGCCGGTATGCCGCAAATTG 621
          || ||||||||||| | ||||||||| |  ||  |||||||||  ||||||
Db   1791 GCTGGGTGGTGGCGGCGCGGGTATGCCGGGTATGCCGGCGATCGCTGATAATTGTCAAACTG 1850

Qy    622 GCAACAGAAGACATCTGATGAGCTGAAGCGGCATTCGGCGGACGAGGTTCGGTATATGTCGAA 681
          || |  |||| || || | |||| |||||||||| |  |  | |  | || | || || ||
Db   1851 TTCTTAGCCGGCCAAGGTGAAGCTGGTGCAGCCTTTGGTAATGCTCGGCGTTTATATAGAA 1910

Qy    682 CGTGCTGATTAACCCCAGCACATTGAAGTGCAGACTCCTTGGCCATCGCACTCGAGAA 741
          ||||||||||||  | || || | || ||| |  |||  | |  | || ||| || ||
Db   1911 AAATTTATTGAACGTCGCGCCACAATTTCAAATTTTGGCTGTGATAATTACGGCAAT 1970

Qy    742 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTGTCACCAAAAAGTTGTCGAA 801
          ||||| ||| |||  |||||| ||  |||| | |||||||| || |  |||||| | ||
Db   1971 GTGATTCACTTGGGTGAGAGGATTGTGAACGTGCGTCGTAACCAAAAGTTACTAGAGA 2030

Qy    802 ATTGCGGCAGCACAGCATTTGGATCCAGAACTGCGTGATCGGCATTGTGCCGATGCAGTA 861
          |||| | ||| |  | || || | ||| || ||  |||   |  || ||| ||  ||||
Db   2031 GAAGCCCCCAGCCCAGCCCAGCTTGGACTCAGAACCTAAGGGAAAAAATGGGACAAGCGGCGGTG 2090

Qy    862 AAGTTCTGCCGTCCATTGGTTACCAGGGGCCGGGAACCGTGGAATTCTTGGTCGATGAA 921
          |||  |||  |   |||  ||  | |||   ||||    ||||  | |  | || || ||
Db   2091 AAAGCGGCTCAGTTTATCAATTATCAATACCGCGGGGCAGGTACTACTCGAGTGTTTGCTGAGTAGA 2150

Qy    922 AAGGGCAACCACGTTTCATCGAAATGAACACCGTATCCGGATTCAAGTAGAGACAATCCCGTAACT 981
          ||||| |||| || |||  |   |   |   | | || | |||||| |  ||||| |||  |||
Db   2151 TCCGGTCAGTTTACTTTATGGAGTGAATGAACACCCGGATTCAAGTGGTGCAACC 2210

Qy    982 GAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCCGTCGGTGCACC 1041
          |||| |||  ||||||  | || ||| || ||||| ||  |||  | ||||
Db   2211 GAGATGGTTACTGACGTGGATTTATTGGTTGAGCAAATCAGAATTGCCAAGGGGAA-- 2268

Qy   1042 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGTGCAGCACTGCAGTGC 1101
          ||| |   | |  ||| | || | || || |||| | | || | || || |  |||||
Db   2269 ------GACTTAGACTCAAGACCAAGTAGTTTACGGCGGTCATGCGGTCGATGATGAATGT 2321

Qy   1102 CGCATCACCACGGAGATCAAACAACACGGCTTCCGCCAGATACCGGATATCATCACCGCG 1161
          |||||| | ||| | |  ||| |  ||  ||  ||||| |   | || | |||  ||| |
Db   2322 CGCATCAATGCCGAAGACCCCAGACCACGATTTCCGCCCAGCAGGACGGCATTAGCGGT 2381

Qy   1162 TACCGGCTCACCAGGCCGGAGCTGGCGTTCGTCTTGACGGTCCGTGCGTTTCCGTCGCAGCA 1221
          |||||  ||||||| | ||  |||    |||    | | ||   |  |||
```

Db 2382 TATCTTCCCCTGGCGGCCCTGGCGTGCGGATTGACTCCCACGTTTACACGGATTACCAA 2441
Db   729 TTCCGCATTTTTGACAGCTTGAAACTGGGTTAAAGGATGAGCGCTCGCCATTGATGCAGTG 788

Qy 1222 ATCACCGCACACTTTGACTCATCCATGCTGGTGAAAAATGACCTGCCGTGTTCCGACTTTGAA 1281
Qy 2044 CTGGAGACCAACACCGGTAGCCGGAGGTGGCTATGGCTATTCTGTGATCTCTCTGAT 2103

Db 2442 ATTCCGCCCTACTACGATTCCTTAATTGCTAAATTGATCGTTTGGGGCCCTGATCGCGCCT 2501
Db   789 CGGCGAGTCAGGCA---AGCTTGCCGAAGCGGCGGATCGTCGTATACAGGGACATTCTCGAT 845

Qy 1282 ACTGCTGTGTGCTGCTGCGTGCGACGGCCGGTTGGCTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1341
Qy 2104 CCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGGATCGTCAAG 2163

Db 2502 ACTGCTATTAACCGCATGAAACGCGCCCTCAGGGAATGCGCCATCACTGGATTACCTACA 2561
Db   846 CCGAACAGAAGCAAGTATAATCTTGAATATTATACGTCAATGGCAAAAAGAGCTTGAAGCT 905

Qy 1342 AACATTGGTTT 1352
Qy 2164 TCTGGGCCTCACATTCTGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTA 2223

Db 2562 ACCATTGGGGT 2572
Db   906 GCGGGGCGCGCCATATTCTCGGCCATTAAAGAGATATGCCGGCCTGCTGCACCTCAAGCGGCA 965

RESULT 15
US-09-974-300-2

Qy 2224 ACCAAGCTGGTCACCGCACTGCGCCGTGAATTCGATCTGCCAGTGCACGTGCACACCAC 2283
; Sequence 2, Application US/09974300

Db   966 TACGAACTGGTGTCGGCTTGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCT 1025
; Patent No. US2002014672lA1

; GENERAL INFORMATION:
Qy 2284 GACACTGCGGGTGGCCAGTTGCTACCTACTTTTATGTATCGCAAGCGATAGAAGCAGGGCGTCGACATC 2343

; APPLICANT: Clausen, Ib Groth
Db 1026 GACACGAGCGGCAACGTATTTTTATGTATCGCAAGCGATAGAAGCAGGGCGTCGACATC 1085

; APPLICANT: Berka, Randy M.
Qy 2344 GTTGACGGTGCTTCCGCACCACTGTCTGGCCACCACTCCCAGCCATCCCTGTCTCGCCATT 2403

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
Db 1086 GTCGACGTAGCGGTCAGCTCGATGGCGGGTCTGACATCACAGCCAAGCGCAAGCTCGCTT 1145

; TITLE OF INVENTION: Expression
Qy 2404 GTTGCTGCCATTCGGCGCACACCCGTCGCGATACCGGTTGAGCCTGCAGGCTGTTTCTGAC 2463

; FILE REFERENCE: 10085.500-US
Db 1146 TACCATGCGCTTGAAGGAGATAAACGCCGTCCCGCAGTTCAATGTCGATGCGGAAGAAATAA 1205

; CURRENT APPLICATION NUMBER: US/09/974,300
Qy 2464 CTCGAGCCGTACTGGGAAGCTGTGCCGCGACTGTACCTGAGTCTGGAACCCCA 2523

; CURRENT FILING DATE: 2001-10-05
Db 1206 TAGTCTCAAAAAAAAGTCTGTCAGAAAATATTACAGCGAGTTTGAGAGCGGCATGATT 1265

; PRIOR APPLICATION NUMBER: 09/680,598
Qy 2524 GGCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTCGT 2583

; PRIOR FILING DATE: 2000-10-06
Db 1266 GCGCCTCATACCGCGAAATTACAGCATGAAATGCCTGGCGGCCAATACAGCAACCTTCAG 1325

; PRIOR APPLICATION NUMBER: 60/279,526
Qy 2584 GCACAGGCCACCGCACTGCGCCTTGCTGATCGCTTCGAGCTTCATCGAGACAACTACGCA 2643

; PRIOR FILING DATE: 2001-03-27
Db 1326 CAGCAGGCTAAGGGAGTCGGGCCTGGGACGCGCCAACCAAGTCACCCCATCCTCCAAGGTTGTTGGC 1385

; NUMBER OF SEQ ID NOS: 8481

; SOFTWARE: FastSEQ for Windows Version 4.0
Qy 2644 GCCGTAATGAATGAGATGCGGGGGACACCACCCATCCTCCAAGGTTGTCGGG 2703

SEQ ID NO 2
Db 1386 CGGGTCAACCACCTGTTCGGGGACATCGTAAAGGTTACGCCGTCATCTAAAGTCGTCGGG 1445

; LENGTH: 2238
Qy 2704 GACCTCCGCACTCCACCTGGTTGGTGCGGGTGAGATCCAGACTTTGCTGCAGACCCCA 2763

; TYPE: DNA
Db 1446 GATATGGCGGTTTACATGGTGCAAAATAACCTGACGGAAGATGATATTTACGAAAAGGGGA 1505

; ORGANISM: Bacillus licheniformis
Qy 2764 CAAAAGTACGACATCCCAGACTTCGTCATCGCGTTCCTGCGCGAGCTTGGTAACCCT 2823

US-09-974-300-2
Db 1506 GAATCTCTCGATTTTCCTGACTCGGTTGTCGAGCTTTTTAAAGGGTATATCGGTCAGCGCT 1565

Query Match   8.2%;   Score 286; DB 10; Length 2238;
Qy 2824 CCAGGTGCGTGGCCAGAACCACTGCGCACCGCGAAGGCCGCTCCCGAAGGCAAG 2883

Best Local Similarity 48.8%; Pred. No. 4.9e-78;
Db 1566 CACGGCCGGATTCCCTGGAGAAATTGCAAAAGTCAAAAGCTGATTTTAAAAGGGCCAGGAGGCCGATTACA 1625

Matches 939; Conservative   0; Mismatches 955;   Indels   30;   Gaps   30;
Qy 2884 GCACCTCTGACGGAAGTTCCTGAGGAGAGACAGGCGCACCTCGACGGCTGATGATTCCAAG 2943

Db 1626 GTACGCCCTGGAGAACTGCTTGAACCGGTTTCATTGACGTATAAAAGCGGAATTCTTG 1685

Qy 1564 GATCTGCCACTGCCACGCCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTT 1623
Qy 2944 GAACGTC-------GCAACAGCCTCAACCGCCTGCTGTTC-------CCGAAG 2982

Db   309 GATCAGCCAATCGCAAGCGGAACAAAACAGATTCTTGATGAACGCGAGCCGAAGGGGCTC 368
Db 1686 GAAAAGCATGGTATGGAGCTTTCCGACCAGGATGCGTGTTGCATATGCGCTTTATCCGAAA 1745

Qy 1624 GCTCGTGATCTCCGTGAGCGAGGACGACGCACTGCAGTTACTGATACCACCTTCCGCGATGCA 1683
Qy 2983 CCAACCGAAGAGTTCCTGCGACCAGTCTGCCAGCTCGCGCGGTATACAAAC 3042

Db   369 GTCAAATGGGGTGAAAGATCAAGAGGAGAGTGCTCCCTCACTGATACGACGTTCCCGGACGCC 428
Db 1746 GTGTTTACCGGAGTATGTCAAGACAGCCCGAGCGCCGGAACTGTACGGGGATATCTCGTGTCTTGATACA 1805

Qy 1684 CACCAGTGCTTTGCTGCGCACCGGAGTCCGCTCATTCGCACTGCGAGCCTGCGGCGAGAGGCC 1743
Qy 1924 ATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCGCAACAACCGGTGGATACACCCCCG 1983

Db   429 CATCAGTCATTGCTTGCGCAGAAGAGTCAGAACGCATGACCTGAAAAAAATCGCCAATCCG 488
Db   609 ATTCCGAATACGACGTGTTCAAATGCTTTGAAGATCTTGAGGAAGCTGCCCAGCTCGCGCGGTATACAAAC 668

Qy 1744 GTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCGCTGGGGCGGGCCGACCTACGAT 1803
Qy 1924 GTGTTTACCGGAGTATGTCAAGACAGCCCGAGCGCCGGAACTGTACGGGGATATCTCGTGTCTTGATACA 1983

Db   489 ACGGCTGCGCTCTCGGCCAGAGCTTTCAGTCTTGAAATGTGGGGCGGCGGGCCCCACATTGAT 548
Db   669 TACCCTGACAACCTGATTAAAAAGTTCGTCAGCGAATCGCGCGCGGCCCGAAATCGATGTA 728

Qy 1804 GTGGCGATGCGTTTCCTCTTTGAGATCGTTTGGGGGATCGTGGGGACAGGCTGACGCTGCGCGAGGCG 1863
Qy 1984 TTCCGCGATCTTGACGCGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTC 2043

Db   549 GTCGCGTTACCGTTTCTTCTTGAAAGAAGAGATCCGTGGAAAAAGGCTGCAGGAGCTGCGGAAGGAA 608
Db 1806 CCAACCTTCCCTCTACGGCATGAACTGCAGTCTCCAGCAATCGCGACGCAGTC 1862

Qy 1864 ATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCCGCAACAACCGGTGGATATCACCCCCG 1923
Db 2983 CCAACCGAAGAGTTCCTGCGACCAGTCTGCCAGCTCGCGCGGTATACAAAC 3042

```
Qy   3103  CGCACCCCACTGCTTGTTCGCCTGAGATCTTCGCCTGAGATCTCTGAGCCAGACGATAAGGGTATGCGC  3162
                 --||||||||||-------||||---||||||-----------------||--------
Db   1863  GGGAAAACGCTGATCGTCAAACTTGTCTCTATCGGGGAACCGCGTCCGATGCCGACGAGA  1922

Qy   3163  AATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTT  3222
                 ||-----||||||||||-------||--||---|||-----------||-
Db   1923  GTCGTTTATTTGAACTGAACGCGGCCAGCCTCGCGAAGTGGTCATTAAAGACGAAAGCATT  1982

Qy   3223  GAGTCTGTCACCGCAACCGCAGAAAAGGCAGATTCCTCCAACAGGGCCATGTGTTGCTGCA  3282
                 -|--------------------|------||||---||||||---------|||--
Db   1983  AAATCGTCCGTTCATCAAAAGGTGAAAGCCGACCGTTCAAACCCGAATCATATCGCGGCA  2042

Qy   3283  CCATTCGCTGG---TGTTGTCACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTGGA  3339
                 --||||---||-------||---||||--|||||----|||---||||---|||---
Db   2043  TCTATGCCGGGAACGGTAATCAAGCTTCTTGTAAGCAAGGTGACAAGTGAAGAAGAGGGC  2102

Qy   3340  GATGCAGTCCCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGAC  3399
                 |||-------||-||||||||--|-|||||-||||||------||----|---||--
Db   2103  GATCATTTGATGATCAATGAAGCGATGAAAATGAAACGACCGTTCAGGCTCCGTTTTCA  2162

Qy   3400  GGCAAGATTGAACCGCGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATC  3459
                 ||||||---||||||----||||||||----||||||--------|||||------
Db   2163  GGAACTGTTGAAAACATTCACGTTACAAACGGAGAAGCCATTCAAACCGGCGACCTTCTC  2222

Qy   3460  GTCG  3463
           ||--
Db   2223  ATTG  2226
```

Search completed: April 5, 2003, 12:50:28
Job time : 366 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2003, 04:31:45 ; Search time 5238 Seconds
(without alignments)
10741.337 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtgactgctatcacccttgg........tgatcgtcgtgttcctaa 3474

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

	Result No.	Score	% Match	Query Length	DB	ID	Description
c	1	237.4	6.8	888	17	CNS074FW	AL428754 clone BA0
c	2	234.4	6.7	903	14	BQ881271	BQ881271 AGENCOURT
c	3	221.8	6.4	859	17	AQ875053	AQ875053 V120F6 mT
	4	220.2	6.3	637	9	BF251052	BF251052 EST418309
	5	218	6.3	675	2	AI109467	AI109467 GH08678.5
	6	213.2	6.1	744	13	BI103175	BI103175 602889415

	No.	Score	%	Length	DB	ID	ID	Description
	7	207.8	6.0	640	13	BI618593	BI618593	RH49459.5
	8	205	5.9	715	12	BG910468	BG910468	602806142
	9	201.2	5.8	641	13	BJ063722	BJ063722	BJ063722
	10	197.4	5.7	887	12	BF781506	BF781506	602104410
	11	196.6	5.7	515	9	AU202878	AU202878	AU202878
	12	196.2	5.6	949	14	BQ752438	BQ752438	EST633001
	13	195	5.6	666	13	BI605019	BI605019	RH70532.5
	14	193.2	5.6	1030	17	CNS06M6L	AL405091	T3 end of
	15	192.6	5.5	625	9	AI293242	AI293242	GH16378.5
	16	192.4	5.5	508	13	BI621845	BI621845	RH53265.5
	17	189.6	5.5	921	14	BQ959497	BQ959497	AGENCOURT
	18	189.2	5.5	895	12	BF535766	BF535766	602051271
	19	189	5.4	611	10	BE312170	BE312170	601152484
	20	186.4	5.4	860	13	BI872878	BI872878	603397491
	21	186	5.4	915	14	BQ730755	BQ730755	AGENCOURT
	22	185.8	5.3	599	12	BF506511	BF506511	AT09003.5
c	23	184.8	5.3	553	17	AQ501087	AQ501087	V27H5 mTn
	24	184.8	5.3	2403	11	AK007782	AK007782	Mus muscu
	25	182.4	5.3	725	12	BG599744	BG599744	EST504639
	26	182.2	5.2	813	12	BG761035	BG761035	602717545
	27	182	5.2	556	9	AI134619	AI134619	GH12314.5
	28	181.8	5.2	575	9	AI389557	AI389557	GH20767.5
c	29	181.8	5.2	923	14	BQ734080	BQ734080	AGENCOURT
	30	181.2	5.2	610	13	BI082302	BI082302	602877488
	31	181	5.2	643	12	BF499491	BF499491	AT14044.5
c	32	179.6	5.2	780	17	BH770815	BH770815	LLMGtag55
	33	178.6	5.1	854	13	BI146211	BI146211	602911028
	34	178.4	5.1	684	13	BI718188	BI718188	1031024D0
	35	177.6	5.1	632	13	BJ124629	BJ124629	BJ124629
	36	176.8	5.1	609	9	AI526573	AI526573	uj40b10.y
	37	175.6	5.1	616	12	BG710300	BG710300	pg11n.pk0
c	38	174.6	5.0	848	17	AQ875041	AQ875041	V120E5 mT
c	39	173.8	5.0	856	17	CNS0749M	AL428528	clone BA0
	40	173.4	5.0	635	13	BM156619	BM156619	fv57d12.y
	41	173.4	5.0	794	13	BI683931	BI683931	603306761
	42	173.2	5.0	725	13	BI327807	BI327807	602979556
	43	173	5.0	652	12	BG855209	BG855209	1024042B0
	44	169.6	4.9	598	13	BJ125540	BJ125540	BJ125540
	45	168.8	4.9	683	13	BM603525	BM603525	170006870

ALIGNMENTS

RESULT 1
CNS074FW/c
LOCUS   CNS074FW   888 bp   DNA   linear   GSS 07-JUL-2001
DEFINITION clone BAOAB027F04 of library BAOAB from strain CLIB 210 of
            Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL428754
VERSION   AL428754.1 GI:12211948
KEYWORDS   GSS.
SOURCE   Kluyveromyces lactis.
ORGANISM   Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1 (bases 1 to 888)
AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE   20584711
PUBMED   11152876
REFERENCE   2 (bases 1 to 888)
AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 11.

JOURNAL   Kluyveromyces lactis
           FEBS Lett. 487 (1), 66-70 (2000)
MEDLINE   20584721
PUBMED   11152886
REFERENCE 3 (bases 1 to 888)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
           2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
           segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
           yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
           exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
           Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
           lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
           angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
           Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
           5 kb were prepared and both extremities were sequenced. See
           keywords for description of this sequence and for the sequence of
           the other extremity of this insert.
FEATURES   Location/Qualifiers
     source   1..888
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB027F04"
                     /clone_lib="BA0AB"
     misc_feature   complement(<2..>886)
                     /note="similar to Saccharomyces cerevisiae ORF YBR218C [
                     PYC2 ; pyruvate carboxylase 2]
                     similar to Saccharomyces cerevisiae ORF YGL062w [PYC1 ;
                     pyruvate carboxylase 1]"
                     /evidence=not experimental
BASE COUNT   249 a   203 c   186 g   250 t
ORIGIN

Query Match   6.8%;   Score 237.4;   DB 17;   Length 888;
Best Local Similarity   55.7%;   Pred. No. 4.7e-52;
Matches 476;   Conservative   0;   Mismatches 376;   Indels   3;   Gaps

Qy 248 CACCAGTCAAGGCGTACCTGACATGCATGGATGAAATTATCGGTGCAGCTAAAAAGTTAAAG 307
            |||| || | || || | ||| || || || |||||| |||||||||||
Db 855 CTCCAGTTGGTGCTTACTTGGCGATTGATGAAATTATCAACATTGCCAAGTCCCATGGTG 796

Qy 308 CAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGAAAAATGCCCAGCTTGCCCCGCGAGT 367
            || | | || |||||||| |||| || ||||| ||| ||| ||| |||||
Db 795 TGAACTTTATCCATCCTCGGTTACGGGTTCCTATCAGAAAACTTCGAAATTCGCTGAAAAAG 736

Qy 368 GCGCGGAAAAACGGCCATTACTTTTTATTGGCCCAACCCCAGAGTTCTTGATCTCACCGGTG 427
            || || ||| |||| || || ||| | || |||| ||| ||| ||| ||
Db 735 TTGCAGCTTCTCGGTATCCACTTGGGTTGTCTGTTGGTTCGATTCGGTTGGTG 676

Qy 428 ATAAGTCTCGTGCGTAACCGCGCGAAGAAGGCTGGTCTGCCAGTTTTGGCGGAATCCA 487
            ||||| | ||| ||| ||| || | || |||| ||| ||| || |||||| |||
Db 675 ACAAAGTTTCAGCCAGAAACTTGGCTGAAAAGCCTAACGTTCCAGTGGTTCCAGGTACTC 616

Qy 488 C---CCCGAGCAAAAACATCGATGACATCGTTAAAAGCCGCTGAAGGCCAGACTTACCCCA 544
            | |||| || || ||||| | |||||||| | ||| | ||| ||||| |||||||
Db 615 CAGGTCCAATTGAGACTGGAGAAGCACAGCCTTCGTTGATAAAATACCGGATTCCCAG 556

Qy 545 TCTTTGTAAAGGCCAGTTGCCGGTGGCGGACGCGGTATGCGCTTTGTTCTTCACCTG 604
            ||||| || ||| | ||| | || || |||| |||| || ||| | || ||
Db 555 TCATCATTAAGGCCGCCCTTTGGTGGTCGTGGTGGTATGAGAGTGGTGTCCCGTGAAGGTG 496

Qy 605 ATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGGCGGCATTCGGCGACG 664
            || || |||||| ||| | || || |||| | |||||||| |||| || |||| |||||
Db 495 ATGATATCAGTGACGCGTTTCCAACGTGCCAAATCTGAAGCTATCACTGCTTTTGGTAACG 436

Qy 665 GTTCGGTATATGTCGAACGTGCTCTGATTAACCCCCAGCACATTGAAGTGCAGATCCTTG 724
            || ||| | | ||| | || | ||| ||||| ||| ||| ||||| ||| || |
Db 435 GTACTTGTTCATTGAAAGATTCTTGGACAAGCCAAAGCATATTCAATTGTTGG 376

GGATCGCACTGGAGAAGTTGTACACCTTTATGAAGTGACTGCTCACTGCAGCGGTCGTC 784
Qy 725 GCGATCGCACTGGAGAAGTTGTACACCTTTATGAAGTGACTGCTCACTGCAGCGGTCGTC 784
            ||| | | || | ||| ||||||||||| || || || ||| |||| |||| |||
Db 375 CTGATAACTACGGTAACGTTGTTCATCTGTTCGAAAGAGATTGTTCCGTTCAAAGAAGAC 316

Qy 785 ACCAAAAAGTTGTCGAAATTGCGCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCA 844
            | || || || |||||| ||| | |||| ||| || |||| || |||| |||||| |
Db 315 ATCAAAAGGTTGTCGAAGTTGCCCAGCCAGCCAAAACTCTACCTGAAGAAGTGCGTGAGCTA 256

Qy 845 TTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGG 904
            ||||| |||| |||| || ||| || | || ||| || |||||| |||| ||| ||
Db 255 TCTTAACTGATGCGGTCAAATTAGCTAAGGAAGCTGGTTACAGAAACGCTGGTACCGCTG 196

Qy 905 AATTCTTGGTCGATGAGAAAAGGGCAACCACGTTTTCATCGAAATGAACCACGTATCCAGG 964
            |||| | |||| |||| ||| | | ||||| || ||||||| |||||| |||| ||||||
Db 195 AATTCTTAGTTCGATAACCAAACAGACATTACTTTATCGAAATTAACCCAAGAATTCAAG 136

Qy 965 TTGAGCACCGTGACTGAAGAGTCACCGAGTGGACCTGGTGAAGGCGCAGATGCGCT 1024
            |||| | |||| | |||| |||| | ||| |||| | ||| || ||| |||| |||
Db 135 TCGAACATACAATTACCCAAGAAATCACCGGTATCGATATCGTCGCTCAAATCCAAA 76

Qy 1025 TGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACACCACG 1084
            ||| | |||| | ||||| | || |||| |||| |||||| ||||| | ||| |||| |
Db   75 TCGCTGCAGGTGCCACTTTGGAACAATTGGGTTTTGATGCAAGATAAGATTACAACTCGTG 16

Qy 1085 GTGCAGCACTGCAGT 1099
            ||| | |||| |||
Db   15 GTTTCGCCATCCAGT 1

RESULT 2
BQ881271
LOCUS   BQ881271   903 bp   mRNA   linear   EST 16-AUG-2002
DEFINITION AGENCOURT 7978339 Lupski dorsal root ganglion Homo sapiens cDNA
            clone IMAGE:6185002 5', mRNA sequence.
ACCESSION   BQ881271
VERSION   BQ881271.1 GI:22273279
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 903)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13575 row: b column: 11
            High quality sequence stop: 692.
FEATURES   Location/Qualifiers
     source   1..903
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185002"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life

```
VERSION      AQ875053.1  GI:6287297
KEYWORDS     GSS.
SOURCE       Saccharomyces cerevisiae
             baker's yeast.
  ORGANISM   Saccharomyces cerevisiae
             Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
             Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE    1  (bases 1 to 859)
  AUTHORS    Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,
             deStages,S.A., Cheung,K.-H., Sheehan,A., Symoniatis,D., Jansen,R.,
             Umansky,L., Heidman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
             Hager,K., Miller,J., Roeder,G.S. and Snyder,M.
  TITLE      Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
             Gene Disruption
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Kumar A
             Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
             Yale University
             P.O. Box 208103, New Haven, CT 06520-8103, USA
             Tel: 203 432 9949
             Fax: 203 432 6161
             Email: anuj.kumar@yale.edu
             te of mTn-3xHA/lacZ insertion.
             Seq primer: GGCCTTCTTCTTGGAAGTAC
             Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1..859
                     /organism="Saccharomyces cerevisiae"
                     /strain="Y2278 - S288C background, cir(0) rho(0)"
                     /db_xref="taxon:4932"
                     /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
                     /lab_host="E. coli"
                     /note="Vector: pHSS6-Sal; A yeast genomic DNA library
                     without 2 micron or mitochondrial DNA was prepared in
                     pHSS6-Sal; genomic DNA was size-fractionated (DNA of
                     roughly 2-3 kb in length) prior to cloning. This library
                     was subsequently mutagenized with a mTn-3xHA/lacZ
                     minitransposon containing lacZ, URA3, and tet resistance.
                     "
BASE COUNT    213 a    196 c    190 g    256 t    4 others
ORIGIN
```

```
Query Match     6.4%;  Score 221.8;  DB 17;  Length 859;
Best Local Similarity  59.8%;  Pred. No. 6.9e-48;
Matches  406;  Conservative  0;  Mismatches  269;  Indels  4;  Gaps  2;

Qy   680  AACGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCACTGGAG 739
              ||| ||| |||||||||||| |||||||| ||||||||
Db   801  AAAATTCTTGACAAAGCAAAAGCATATTGAAGTTCATTGTGGCCCGATAACCACCGAAA 742

Qy   740  AAGTGTACACCTTTATGAACGTGACTGCTCACTGCGAGCGTCGTCACCAAAAAGTGTCG 799
              |||||||||||||||||| ||||||||||||| |||||
Db   741  CCGTTGTTCATCTTTTCGAAAGAGACTGTTCGGTCAGAGAAA-CACCAAAAGGTTGTCA 683

Qy   800  AAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTGTGCGGATGCAG 859
              |||||||  |||||||| ||||  |||||||| ||||||||
Db   682  AAGTGCGNCCAGCAAAGACTTTACNCCGTGAAGTCGGTGACGCCATTTTGACAGATGCAG 623

Qy   860  TAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATG 919
              |||||||||||| ||||||| |||| ||||||||||||
Db   622  TTAAATTGGCCAAAGATGTGGCTACAGAAATGCGGGTACTGCTGCTGAATTCTTGGTTGATA 563

Qy   920  AAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGA 979
              ||||||  ||||||| |||||||||  |||| ||||||||||
Db   562  ACCAAAATAGACACTATTTCATTGAAATTAATCCAAGTGGAACATACCATCA 503

Qy   980  CTGAAGAAGTCACCGAGGTGACCTCGGTCGAAGGGCCGCAGAGTGGCGTTGGCTGGTCGA 1039
              |||||||||||||| ||||||||||||||| |||||
Db   502  CAGAAGAAAATTACCGGTATAGATATTGTGGCGGCCGCTCAGATCCAAATTCAAAATTGCGGCAGGTGCCT 443

Qy  1040  CCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGTGCAGCACTGCAGT 1099
              ||||||||||||| ||||||||| ||||||
Db   442  CTCTACCCCAGCTGGGCCTATTCCAGCAAATTACGACTTCGTGGCTTTGCCATTCAGT 383
```

```
BASE COUNT    172 a    281 c    280 g    170 t
ORIGIN
```

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through life Technologies."

```
Query Match     6.7%;  Score 234.4;  DB 14;  Length 903;
Best Local Similarity  57.3%;  Pred. No. 3e-51;
Matches  465;  Conservative  0;  Mismatches  341;  Indels  6;  Gaps  2;

Qy  1596  CCTGAAGCAGCTTGCCCAGCCGCGTTGCTCCGTGATCGCAGCAGGACGCACTGGC 1655
              ||| ||| |||||| |||||||||||||||| ||||||| |||||
Db    18  CCTGCTGCGAGAGGGGCCTGAGGGCTTTGCTCGAGCTGTGCGGAACCACCCGGGGCTGCT 77

Qy  1656  AGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTC 1715
              ||||||| ||||||  |||||||||||||||| |||||||
Db    78  GCTGATGGACACGACCTTCAGGGAGCACGCCCACCAGTCACTGCTGGCCACTCGTGTCGCGCAC 137

Qy  1716  ATTCGCACTGAAGCCTGCGGCAGAGGCGGTCGCAAAGTGACTCCTGAGCTTTTGTCCGT 1775
              ||||||||||| |||||  ||||||| || ||||||||
Db   138  CCAAGATCTCAAAAGATGCCCCCTATGTGTGCCCACAACTTCAGCAAGCTCTTCAGCAT 197

Qy  1776  GGAGGCCTCGGGGCCGGCGGCGACCTACGATGTGGCGATGCGTTCCTCTTGGAGGATCCGTG 1835
              |||||||||| ||||||||||||| |||| |||||||
Db   198  GGAGAACTGGGAGGGAGGACGACCGTTGACGTCGGAGGCCATTGAGTGACTCCCTG 257

Qy  1836  GGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCG 1895
              |||||  |||||||||||||| ||||||||
Db   258  GCGCGGGCTGCAGGAGCTCCGGAGCTCATCATCCCAACATCCCTTTCCAGATGCTGCTGCG 317

Qy  1896  CGGCCGCAACACCGTGGGATACACCCCGTACCCCAGACTCCGTTGCCGCGCGTTTGTTAA 1955
              ||||||||||| ||||||||||| |||||||||||||
Db   318  GGGGGCCAATGCTGTGGGCTACACCAACATCCAGACAACAGCTGGTTCTCAAGTTCTGTGA 377

Qy  1956  GGAAGCTGCCAGCTCCGGCGTGACATCTTCCGCATCGCCTTAACGACGTCTC 2015
              ||| ||| ||||||| |||||||| |||||||
Db   378  AGTGGCCAAAGAGAATGGCATGGATGTCTTCCGTGTGTTGACTCCCTCAACTACTTGCC 437

Qy  2016  CCAGATGCGTCCAGCAATCGACGCAGTCCTGAGACCAACACCGCGGGGTAGCCGAGGTGGC 2075
              |||| ||||||  ||||||||||
Db   438  CAACATGCTGCTGGGCATGCAGGCCG---CAGGAAGTGCCGGAGGCGTGGGAGGCTGC 494

Qy  2076  TATGGCTTATTCTGTGATCTCTGATCCCAAATGAAAAGCTCTACACCCTGGATTACTA 2135
              |||||||  ||||| ||||||||| |||||
Db   495  CATCTCATACACGGGCGACGTGGCCGACCCCAGCCCGACCAGTACTCACTGCAGTACTA 554

Qy  2136  CCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCCGTCACATTCTGGCCATTAAGGATAT 2195
              |||||| ||||| ||||| |||||||| ||||||
Db   555  CATGGGCGTTGGCCGAAGAGCTGGTGCCGAGCTGGCACCCCATCCTGTGCATCAAGGACAT 614

Qy  2196  GGCTGGTCTGCTTCGCCCAGCTGCGTAACCAAGTGGTCACCGCACTGCG---CCGTGA 2252
              |||||||||  ||||||| |||||||| |||||
Db   615  GGCCGGGCTGCTGCTGGAGCTGCTTGTGCTGGTCAGCTCCCTCCGGGACCGGCTT 674

Qy  2253  ATTCGGATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGTTGGCTACCTA 2312
              ||||||||||||| |||||||||||||||
Db   675  CCCCGACCTTCCCACTGCCACATCCACCACCAACGTGCAGGGGACGAGCCTGGCAGCCAT 734

Qy  2313  CTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGTTGACGGTGCTTCCGCACCACTGTCTGG 2372
              |||| ||||| |||  ||| |||||||||
Db   735  GCTGCGCCTGTGCCCAGGCTGCGAGCTGAGTGTGGATGTGGCAGCTGCAGTCCATGTCTGG 794

Qy  2373  CACCACCTCCCAGCATCCCTGTCTGCCATTG 2404
              |||| ||||||  ||||| ||||
Db   795  GATGACTTCACAGCCCAGCATGGGGGGCCCCTG 826
```

```
RESULT 3
AQ875053/c
LOCUS       AQ875053   V120F6 mTn-3xHA/lacZ Insertion Library,
DEFINITION  V120F6 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
            cerevisiae genomic 5', DNA sequence.
ACCESSION   AQ875053
```

AQ875053   859 bp   DNA   linear   GSS 08-NOV-1999

Qy 1100 GCCGCATCCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCG 1159
         ||||| || ||||||| | |||| | ||| ||||||||| |||||| ||
Db 382 GCCGTATTACCACCGGAGAGACCCTGCTAAGAACTTCCAACCAGATACCGGTAGAATAGAAG 323

Qy 1160 CGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCGTCTTGACGGTG---CAGCTCAGCTCCGTG 1216
         || ||||| ||||||||||| |||| |||| ||||| | || | |||| |||||
Db 322 TGTACCGGTTCTGCAGGTGGTAATGGTGTTAGACTGGATGGTGTAACGCCTATGCAGGAA 263

Qy 1217 GCGAAATACCGCACACTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACT 1276
         ||||| ||||||||| | | || || ||| | ||||| || ||| |||||
Db 262 CAATAATCTCACCTCATTACGACTCAAATGCTGGTCAAATGCTTCATGCTCCGGTTCCACCT 203

Qy 1277 TTGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTG 1336
         ||||| |||| ||| |||||| | |||||||||| ||||||
Db 202 ACGAAATCGTTCGTAGAAAAATGATTCGTGCATTAATCGAGTTCAGATTAGAGGTGTCA 143

Qy 1337 CAACCAACATTGGTTTCTT 1355
         ||||||||| |||| ||
Db 142 AGACCAAACATTCCCTTCCT 124

RESULT 4
BF251052
LOCUS   BF251052   637 bp   mRNA   linear   EST 15-NOV-2001
DEFINITION Coccidioides immitis spherule cDNA library Coccidioides
            immitis cDNA clone CIAAB85 5' sequence, mRNA sequence.
ACCESSION   BF251052
VERSION   BF251052.1 GI:16931195
KEYWORDS   EST.
SOURCE   Coccidioides immitis.
ORGANISM   Coccidioides immitis
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1 (bases 1 to 637)
AUTHORS   Gardner,M.J. and Kirkland,T.
TITLE   Generation of ESTs from Coccidioides immitis spherule cDNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Malcolm J. Gardner
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: gardner@tigr.org.
FEATURES   Location/Qualifiers
     source   1..637
                     /organism="Coccidioides immitis"
                     /db_xref="taxon:5501"
                     /clone="CIAAB85"
                     /clone_lib="Coccidioides immitis spherule cDNA library"
                     /dev_stage="spherule"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT   123 a   261 c   163 g   90 t
ORIGIN

Query Match   6.3%; Score 220.2; DB 12; Length 637;
Best Local Similarity   60.4%; Pred. No. 1.5e-47;
Matches   380; Conservative   0; Mismatches 248; Indels   1; Gaps   1;

Qy 538 TACCCCATCTTTGTAAAGCAGTTGCCGGTGGTGGCGGACGCCGGTATGCCGTTTGTTTCT 597
         ||||||||| |||| |||| |||
Db   7 TTCCCCATCATCATCAAGGCCGGCCGCTTCGGCCGGCGGCCATGCCGTGTCGTCCGC 66

Qy 598 TCACCTGATGAGCTCCGCAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTC 657
         | | |||||| || || |||||||
Db   67 GACCAGGAGTCCCTCCGCGACTCCTTCGAGACGTCTGATTAACCCCAGCACATTGAAGTGCAG 717

Qy 658 GGCGACGGTCGACCCGTCTTCGTGAGGCGTATGTCGAAGCGTCGTGATTAACGTCGTG 717
         ||||| |||| || |||||
Db 127 GGCAACGGACCACCGGTCTTCCTCGACAGGCCAAGCACATCGAGTGCCAG 186

Qy 718 ATCCTTGGCACTGCGACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAG 777
         || ||||| |||||| | | | | | | ||
Db 187 CTGCTCCGGCGACAACCACGGCAACGTCGTCCACCTCTACGACGTGACTGCAGCGTCCAG 246

Qy 778 CGTCGTCACCAAAAAGTTGTCGAAATTGCGCAGCACAGCATTGATCCAGAACTGCGT 837
         ||||| ||||| || ||||||
Db 247 CGCCGGCCACCAGCAGAAGGTCGTCGAGCTGCAGCTCCCCGGGCGCCAAGGGGCGCCGC 306

Qy 838 GATCGCATTTGTGCGGATGCAGTAGAGGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGA 897
         || | | | | |||||
Db 307 GACAGCATCCTCGCCACGCGTCAGCTGGCGCTCGACGCTCCGTCAGCTACCGGACATCGCGGGC 366

Qy 898 ACCGTGGAATTCTTGGTGCTGATGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGT 957
         ||||| ||| |||| || ||||||||
Db 367 ACCGCCCAGTTCCTCGTCGACCAGCTCCAGCTCAACCGCTACTACTGGATCCATCCCCCGC 426

Qy 958 ATCCAGGTTGAGCACCACCGTGACTGAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCAG 1017
         || |||||| | || || |||||| |||||
Db 427 ATCCAGGTCGAGCCACACCATCACCGAGGAGATCACCGGAGATCCGTCCGCCCCAG 486

Qy 1018 ATGCGCTTGGCTGCTGTCAACCTTGAAGGAATTGGGTTCGCGCCTCACCCAGGACCGCATCTCC 1077
         || |||||||| ||| ||| | |||| || ||||| ||
Db 487 ATCCAGATCGGCCGCGCCACCCTCGGAGCAGCTCGACCGCCATCACCACCGGAAGATTAAGATCAAG 546

Qy 1078 ACCCACCGTGCAGCACTGCAGTGCCGGATCATCACCACGGAAGAT-CCAAACAACGGCTTCCG 1136
         ||||| | |||||| ||| ||| | |||
Db 547 ACCAGAGGCGGTTCGCCATCCAGTGCCGGATCACCATCACCACCGCCGAAGATCCCACAAGGGGCTTCCA 606

Qy 1137 CCCAGATACCGGAACTATCACCGCGTACC 1165
         ||||||| | | || |||||
Db 607 GCCCGACACCGGTAAGATCGGAGTCTACC 635

RESULT 5
AI109467
LOCUS   AI109467   675 bp   mRNA   linear   EST 19-APR-2001
DEFINITION GH08678.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH08678 5prime, mRNA sequence.
ACCESSION   AI109467
VERSION   AI109467.1 GI:3477791
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 675)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE   BDGP/HHMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 86 row: G column: 6
            High quality sequence stop: 588.
FEATURES   Location/Qualifiers
     source   1..675
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH08678"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2. Plasmid cDNA library."
BASE COUNT   142 a   214 c   193 g   126 t
ORIGIN

```
Query Match              6.3%;  Score 218;  DB 9;  Length 675;
Best Local Similarity  59.0%;  Pred. No. 6.1e-47;
Matches 393;  Conservative 0;  Mismatches 270;  Indels 3;  Gaps 1;

Qy 1572 ACTGCCACGCGGTTCCCGTGACCGGCTGAGCAGCAGCTGGCCCAGCCGCGGTTGCTCGTGA 1631
          ||| ||  ||||||  ||  |||  ||||  ||  |||  |||
Db    8 ACCACCGAAGGGACTCCGCGAGGTCCTGGTGTGCGAAGGTCCGGAGGCCTTCGCCAAGGA 67

Qy 1632 TCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACCAGTC 1691
Db   68 GGTGCGCAACCGTAAGGAACTGCTACTCATGGACACACCTTCCGCGATGCCCACCAGTC 127

Qy 1692 TTTGTTGCGACCCAGTCGCGCTCATTCCGCCACTGAAGCCTGCGGGGCGGCCGTGCCAAA 1751
Db  128 GCTGCTGGCCACCCGCGTGCGTTCCCAGACTCGCTGAGATCTCCCCCTACGTGACGCA 187

Qy 1752 GCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGCGGCGGCGACCTACGATGTGGCGAT 1811
Db  188 CAAGTTCAACAACCGTATTCGCTGGAGAACTGGGGCGGCGAGCCACCTTCGACGTGGCGCT 247

Qy 1812 GCGGTTCCTCTTTGGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAA 1871
Db  248 GCGCTTCCTCCGCACGAGTGCCCGTGGGAGCGCCGTGGGAGCGCGATGCGCAAGCGCATCCCGAA 307

Qy 1872 TGTAAACATTCAGATGCGTTCCGGGCCGCAACACCCGTGGGATACACCCGTACCCAGA 1931
Db  308 CATTCCCTTCCAGATGCGTGTTGCCGGGTTTGTTAAGGAAGTGCCAGTCTCCGGTGACATCTTCCGGA 367

Qy 1932 CTCCGTCTGCGCGCCGGTTGTTGCCGGGTTTGTTGAGGAAGTGCCAGTCTCCGGTGACATCTTCCGCAT 1991
Db  368 CAACGTGGTCTACAAGTTCGTGAGCTGGCTGGCGTGTGTGAGCCATGGACATCTTCAGGGT 427

Qy 1992 CTTCGAGACGGGCTTAACGACGTCTCCCAGATCGTCGTCCAGCAATCGACGCAGTCCTGGAGAC 2051
Db  428 GTTCGAGACTGCTCACCTACCTGCCCAACCTGACATCCTCGGCAACCTGGGCATGGAAAGGC 487

Qy 2052 CAACACCGGGTAGCCGGAGGTGGCTATGGCCTATGGCTTATTTCTGGTGATCCTCTCTGATCCAAATGA 2111
Db  488 GGGGGCGGTGGT---GGAGGCTGCCATCTCCTATACCGAGAGGACGTCAAGTCTTCGGCGC 544

Qy 2112 AAAGCTCTACACCCTGGATTACTACTAAAGATGCGCAGGAGATCGACGAGGAGGAGATCGTCAAGTCTGGCGC 2171
Db  545 CACCAAGTATGATCTGAAATACTACACTACTACAACCTTCCTGGTCTGCTTCGCGGGCGGCGAC 604

Qy 2172 TCACATTCTGGCCATTAAGGATATGCTGGTCTGCTTCGCTGCTTCGCCCAGCTTCT 2231
Db  605 CCACGTGCTCTGCATTAAGGATATGCTGCTGTCGTCTGTCAAGCCCGAATCCGCCAGCTTCT 664

Qy 2232 GGTCAC 2237
Db  665 GATCAC 670


RESULT 6
BI103175
LOCUS       BI103175      744 bp    mRNA    linear    EST    26-JUN-2001
DEFINITION  602889415F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044779
            5', mRNA sequence.
ACCESSION   BI103175
VERSION     BI103175.1    GI:11454068
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 744)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/,
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1122 row: m column: 04
            High quality sequence stop: 737.
FEATURES             Location/Qualifiers
     source          1..744
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5044779"
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT          158 a     218 c     218 g     150 t
ORIGIN

Query Match              6.1%;  Score 213.2;  DB 13;  Length 744;
Best Local Similarity  58.4%;  Pred. No. 1.3e-45;
Matches 429;  Conservative 0;  Mismatches 298;  Indels 7;  Gaps 3;

Qy 1618 GCGTTGCTCGTGATCTCCGTGACTCCGTGAGCAGGAGGACGCACTGGCAGTTACTGATACCACCTTCCGC 1677
Db   15 GGGCTTGCCCCAGCGTGCGGAATCACCACGGGCTGCGTCTGTGATGGACACAACCTTCCGG 74

Qy 1678 GATGCACCAGTCTTTGCTGCGACCCGAGTCCGCGCTCATTCCGCACTGAAGCCTGCGCGGCA 1737
Db   75 GATGCCCACCAGTCACTACTGGCCACTAGAGTGCGGCACACATGATCTCAA-AAAGATGCG 133

Qy 1738 GAGGCCGTCGGAAAGTGACTCCTGAGCTTTTGTCCGTGGAGGCCGGCGGCGGCGACC 1797
Db  134 CCCTATGTTGCCCCACAACTTCAAACAGCTCTTGAGGATCGTCTTCAGCATCGGGACAACTGGGGAGGCGGCCACG 193

Qy 1798 TACGATGTGGCGATGCCGATGCGTTCCTCTTTGAGGATCCGTGGGACAGGCTCGCGC 1857
Db  194 TTCGACGTGGCCATGGCGTTCCTGTACGAGTGCGCCCTCGGGCCGGCTCCAGGAGCTCCGG 253

Qy 1858 GAGGGCGATGCGAATGTAAACATTCAGATGCGTTCGGCGCGGCCGCCAACAACCTGGGATAC 1917
Db  254 GAGCTTATCCGAACATCCCGTTCCAGATGCTACTGAGGGGGGCCAATGCTGTGGGCTAC 313

Qy 1918 ACCCCGTACCCAGATCCGTCGTCGCGTGCCGCGGTTGTTAAGGAAGCTGCCAGTCGCCGGGTG 1977
Db  314 ACCAACTACCCCTGACAACGTGGTCTTCAAGTTCGTGAGGTGGCCAAAGAGGAAATGGTATG 373

Qy 1978 GACATCTTCCGCATCTTCGACGCGTTGCGACGTTCCCCAACTACTTGCCAAACATGCTGCTGGGCATGGAA 2037
Db  374 GACGTCTTCCGCAGTCTTTGACTCCCCAACTACTTGCCAAACATGCTGCTGGGCATGGAA 433

Qy 2038 GCAGTCCTGGAGACCAACACCGGTAGCCGAGGTGGCTATGGCCTATGGCTTATTCTGGTGATCTC 2097
Db  434 GCAG---CAGGCAGTGCTGGGGGTGTGGGAGGCTGCCATCTCATCAAAGATGGCAGGAGGACGTG 490

Qy 2098 TCTGATCCAAATGAAAAGCTCATACACCCTGGATTACTACCTAAAGATGGCAGGAGGAGGATC 2157
Db  491 GCTGACCCTAGTCGCACTAAATACTCACTGGAGTACTACACTGGCTTAGCTGAAGAACTG 550

Qy 2158 GTCAAGTCTGGCGGCTACAGTTCTGGCCATTAAGGATATGCTGGTCTGCTTCGCCCAGCT 2217
Db  551 GTGCGAGCTGGCACTCACATCCTGTGCATTAAGGACAATGGGCGGCGGGCCGTGAAGCCTGCC 610

Qy 2218 GCGGTAACCAAGTCGTCACCGCACTGCG---CCGTGAATTGCATCGCCAGTGCACGTG 2274
Db  611 GCCTGCACCAGTCGTGTCAGCTCCGTCGCTCCCGGACCGATTCCCCGACCTCCCACTGCACATC 670

Qy 2275 CACACCCACCACGACGCGCGGTGCCAGTTGGCTACCTACTTTGCTGCGACCTCCAAGCTGGT 2334
```

```
Db    671  CATACCCATGATACATCAGGGGCAGGTGTGGCAGCCATGTCTGGCCTGTGCACAAGCAGGG  730

Qy   2335  GCAGATGCTGTTGA  2348
                ||||| || ||
Db    731  GCTGATGTCGTGGA  744


RESULT 7
BI618593
LOCUS       BI618593                640 bp    mRNA        linear   EST 07-SEP-2001
DEFINITION  RH49459.5prime RH Drosophila melanogaster normalized Head pF1c-1
            Drosophila melanogaster cDNA clone RH49459 5 similar to
            BCDNA:GH06348: FBan0001516 GO:[pyruvate carboxylase (GO:0004736)]
            located on: 2R 46B14-46C1; 08/19/2001, mRNA sequence.
ACCESSION   BI618593
VERSION     BI618593.1  GI:15514118
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE       BDGP/HHMI RH Drosophila EST Project
JOURNAL     Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003831: arm:2R [4876906,5137815]
            estimated-cyto:46B13-46E4; 08/19/2001
            Plate: RH.494  row: E  column: 11
            High quality sequence stop: 543.
FEATURES             Location/Qualifiers
     source          1..640
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH49459"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pF1c-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pF1c; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT    138 a   185 c   193 g   124 t
ORIGIN

Query Match        6.0%;  Score 207.8; DB 13; Length 640;
Best Local Similarity   59.1%;  Pred. No. 3.2e-44;
Matches  375;  Conservative  0;  Mismatches  257;  Indels  3;  Gaps  1;

Qy    643  GCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAG  702
               |||||||| ||||||| ||||| | ||| |||||||  |   || || ||
Db      1  GCCAAGGCCCGCTTTGCAATGGTGCGATGTTCATTGAGAAGTTCATCGAGCGACCGCGT  60

Qy    703  CACATTGAAGCGCAGATCCTTGGCGATCCGCACTGGACTGAGAGAGAGTTGTACACCTTTATGAACGT  762
               ||||| ||||| ||  ||  || ||||  ||||| ||| |||| || ||||| |||| |
Db     61  CACATTGAGGTTCAACTGCGTGGAGACAACAAGGCCGGGGAATGTAGTGCATCTGTACGAACGT  120

Qy    763  GACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAACATTGCGGCCCAGCACAGCATTTG  822
               |||||||||| ||| ||| || | | ||| || || |||| ||  ||||||||||||| |
Db    121  GACTGCTCCGTGCAGCGTCGCCACCAGAAGAGGTGGTGGAGATCGCCCCCAGCGCCACGCCCTG  180
```

```
Db    823  GATCCAGAACTGCGTGATCGCATTGTGCGGATGCAGTCAGTAAAGTTCTGCCGCTCCATTGGT  882
               ||| |||  || | || | || || ||    ||  || |||  ||| ||||| ||   | |
Db    181  CCCATCGAAATCCGGGACAAGATGACCGGAGCAGCAGTGCGCTTGGCCGCCCATGTGGGGC  240

Qy    883  TACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTTCATC  942
               || ||||| || |||||| ||||| || |||  | ||||  || ||||  || ||||||
Db    241  TATGAAAACGCCGGAACCGTGGAGTTCTTTGCGACGAGTCGCAACTTCTTACTTCATC  300

Qy    943  GAAATGACCCCAGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTCGAC  1002
               ||||||| || || ||| ||||| || || ||| ||| ||| || ||| ||| ||||
Db    301  GAAGTGAACGCCGTCTCACAAGTGGAGCACACGGTCACCGAGGAGATAACAGGCATCGAT  360

Qy   1003  CTGGTGAAGCGCCAGATGCGCTTGGCTGCTGCTGGTGCACCTTGAACCTGAAGGAATTGGGTCTGACC  1062
               ||| || | || ||| ||| ||||| |||| |||  ||  |||||| |||  |||||| ||| ||| |
Db    361  TTGGTGCAGTGCAAATCCGGCTGGCCGAGGGCATGACCTACCCGAATTGGGCTCACACG  420

Qy   1063  CAAGATAAGATCAAGACCCAGGTGCAGCACTGCAGTGCGCCATCACCACGGAAGATCCA  1122
               |||| ||  || ||| | ||| | |||  | ||| || ||| | || ||||  ||| ||
Db    421  CAGGACAAGATCTGCGCCGTGCTATGCCATTCAGTGCCGTGTGACCACCGAAGATCCG  480

Qy   1123  AACAACCGCTTCCGCCCAGATACCGGAACTATCACCCGTACCGCTCACCAGGCGGAGCT  1182
               || ||  | |||| |||  ||| ||  |||| ||||  ||  ||||| || || |||
Db    481  GCCAACGATTTTCAGCCCAACACCCGGTCTTCCGATCTGGAGGGGTATG  540

Qy   1183  GGCGTTCGTCTTGTGACGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGAC  1239
               ||||||| |  || |||||||   | |||||| ||  || ||||| ||| ||||| || |||
Db    541  GGCATTAGGTTGGACAGTGCGTGCTCCGCCTACGCCGAGCCATCATTTCGCCATACTACGAC  600

Qy   1240  TCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGA  1274
               |||| | |||| ||| |||  || | || ||| ||
Db    601  TCCCTGCTGGTCAAGGTGATCTCCGCACGCAGCGA  635


RESULT 8
BG910468
LOCUS       BG910468                715 bp    mRNA        linear   EST 05-JUN-2001
DEFINITION  602806142F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4938203
            5', mRNA sequence.
ACCESSION   BG910468
VERSION     BG910468.1  GI:14290944
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10873  row: d   column: 12
            High quality sequence stop: 715.
FEATURES             Location/Qualifiers
     source          1..715
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4938203"
                     /clone_lib="NCI_CGAP_Brn67"
                     /tissue_type="anaplastic oligodendroglioma with 1p/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
```

BASE COUNT   156 a   239 c   210 g   110 t
ORIGIN

Query Match   5.9%;   Score 205;   DB 12;   Length 715;
Best Local Similarity   56.6%;   Pred. No. 1.9e-43;
Matches   401;   Conservative   0;   Mismatches 305;   Indels   3;   Gaps   1;

```
Qy    725 GCGATCGCACTGGAGAGAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTC 784
          |||  ||||| | |||| |||||||||||| |||| |||| |||| ||||||||
Db      1 GGGACCAGTATGGGAACATCCTGCACCTGTACGCGAGACGTGCTCCATCCAGCCGCGGC 60

Qy    785 ACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCA 844
          |||||||||||| |||||| | ||| |||| |||||||| |||| ||||| || ||
Db     61 ACCAGAAGGTGGTCGAGATTGCCCCCGCCGCCCGCAGCTTCGGACTTCGGACTTCGGCG 120

Qy    845 TTTGTGCGGATGCAGTAAAGTTCTGCGCGTCCCATTGGTTACCAGGCGCGGGAACCGTGG 904
          ||| | || |||| | |||| || ||| | |||| |||||||| |||| |||| ||
Db    121 TCACCAGCGACTCTGTGAAACTCGCTAAACAGGTGGGCTACGAACGCAGGCACCGTGG 180

Qy    905 AATTCTTGGTCGATGAGAAAAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATCCAGG 964
          |||  |||||||| ||||| |||| |||| |||||| ||||| |||| |||||| ||||
Db    181 AGTTCCTGGTGGACAGGCACGCACGCAAGCAACTACTTCATCGAGTCAACTCCCGCCTGCAGG 240

Qy    965 TTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCT 1024
          ||||||||||||| |||||||| ||||| ||| ||| ||||||||||| ||||| ||
Db    241 TGGAGCACACGGTCACAGAGGAGATCACCGACGTAGACCTGGTCATGCCTCAGATCCACG 300

Qy   1025 TGGGCTGCTGGTGCAACCTTGAAGGAATTGGGTTCGACCCAGGCCTACCCGGCCTGCGGC 1084
          |||||||||| ||||||||||||| ||||| || ||||||| ||| |||||||
Db    301 TGGGCTGAGGGCAGGAGCCTACCCGACATCACCACG 360

Qy   1085 GTGCAGCACTGCAGTGCCGCGATCGCCGAGATCCAAACAACGGCTTCCGCCCAGATA 1144
          |||||| ||||||| ||||||| ||||| ||||||||||| || |||||
Db    361 GGTGTGCCATCCAGTGCCGGGTACCGGCTCACCAGGCGGAGCTGGGCGTTCGTCTTGACGGTGCAG 420

Qy   1145 CCGGAACTATCACCGGTACCGCTTCACCAGGCGGAGCGGAGGGCATGGGCATCGGCGATAATGCTT 1204
          |||||||||||||| ||| ||| ||||||| |||| |||||| |||| ||||||||
Db    421 CCGGCGCCATTGAGCGTGTTCCGGACGCGGAGCGGAGGGCATGGGCATCGGCGATAATGCTT 480

Qy   1205 CTCAGCTC--GGTGGCGGAAATCACCGCACACTTTGACTCATGCTGGTGAAAATGACCT 1261
          ||||| ||| |||| |||||||||||| |||| ||| || ||||||||||| |||||
Db    481 CCGGCTTCCAAGGAGGCCGTCATCTCGCCCCACTACGACTTCCCTGCTGGTCAAAGTCATTG 540

Qy   1262 GCCGTGGTTCCGACTTTGAAACTGCTGCCACCACCACCGCCGCGGCGCGGTTGGCTGAGTTCA 1321
          ||||||||| || ||||||||||||||| ||| | ||||| |||||| |||||||||
Db    541 CCCACGGCAAAGACCAACCACCACCGCCGCCGGCGCCGCCAAGATGAGCAGGAGGCCCCTTGCCGGAGTTCC 600

Qy   1322 CCGTGTCTCGTGTTGCAACCAACATTGGTTTCTTTGCGTGCGTTGCGTGGGAAGAGAGGACT 1381
          |||| ||||||||||| ||||| |||||| ||||||| ||||||||||| |||
Db    601 GCGTGTCCGAGTGTGAAGACAACAATCGCCTTCCTGCAGAATGTGCTCAACACCACCAGT 660

Qy   1382 TCACTTCCAAGCCAATGATCCGCCACCGATTTATCGGCACAACCACCACACCT 1430
          |||||| ||||||| || ||| ||| |||||||||| |||||||| ||||||
Db    661 TCCTGGCAGGCACTGTGGACACCAGTTCATCGACGCAGAACCCAGAGCT 709
```

RESULT 9
BJ063722
LOCUS   BJ063722   641 bp   mRNA   linear   EST 10-DEC-2001
DEFINITION BJ063722 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL077c10 5', mRNA sequence.
ACCESSION   BJ063722
VERSION   BJ063722.1 GI:17470882
KEYWORDS   EST.
SOURCE   African clawed frog.
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1 (bases 1 to 641)
AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.

TITLE   Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
         Center For Genetic Resource Information
         National Institute of Genetics
         1111 Yata, Mishima, Shizuoka 411-8540, Japan
         Tel: 81-559-81-6856
         Fax: 81-559-81-6855
         Email: tshini@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source   1..641
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL077c10"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett and A.M. Zorn,
                     (Wellcome/CRC Institute)."
BASE COUNT   190 a   142 c   161 g   148 t
ORIGIN

Query Match   5.8%;   Score 201.2;   DB 13;   Length 641;
Best Local Similarity   58.6%;   Pred. No. 1.9e-42;
Matches   369;   Conservative   0;   Mismatches 258;   Indels   3;   Gaps   1;

```
Qy    694 AACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCCGACTCGGACAAGTTGTACACCTT 753
          |    |||| || || ||||||| |||| |||||| || |||||| |||| ||||||
Db      5 AAACCTCGCCACCATTGAAGTTCAGATACTCGTGATAAATGAATGTGTGTACATTTG 64

Qy    754 TATGAACGTGACTGCTCACTGCAGCGTCGTCACCCAAAAAGTTGTCGAAATTGCGCCAGCA 813
          ||||| |||| |||| ||| |||| | ||| ||||| ||||| |||||| |||| |||
Db     65 TATGAAAGAGACTGTTCCGTCCAGAAGAAGCCACCAGAAGGTGGTTGAAATAGCCCCAGCT 124

Qy    814 CAGCATTTGGATCCAGAACTGCGTGATCGACTGCTTGTGCGGATGCAGTAAAGTTCTGCGC 873
          || | ||||| |||||| |||| ||| ||| ||| |||||||| |||| |||||
Db    125 GCACAGCTGGACCTTCAGCTCAGCAGACAGATGACCATGACCTCTGTAAAACTGGCAAAA 184

Qy    874 TCCATTGGTTACCAGGCGCGGGAACCGTGGAGTCTTCGGATCGTGATGAAAACATGGCAGCAT 933
          ||||| |||| |||| |||| || ||||| || |||| |||| ||| || ||| |||
Db    185 CAGGTTGGTTATGAGAATGCCCCAGCGCACTGTGGAGTTCTTGGTTGGATAAACATGGCAGCAT 244

Qy    934 GTTTTCATCGAAATGAACTGACGTTGAGCACACCGTGACTGACTGAAGAAGTCACC 993
          ||||| ||||| ||||| || || |||| |||| || || |||| ||||||| |||
Db    245 TACTTCATAGAAGTCAACTCCAAGTTGAGCATACAGTTGCGTGCTCGTGTGTGTCGCCAAATTACA 304

Qy    994 GAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGGCTGCTGGTGCAACCTTGAAGGAATTG 1053
          ||||||| ||| ||||| ||||| || ||||| |||| ||||| |||| |||||||
Db    305 GATGTTGACCTAGTCCAGTCCAGATCCATGCGCTGGGCACTGCAGAGAGGGAAAAAGCCTTCCTGACCTT 364

Qy   1054 GGTCTGACCCAAGATAAGATCAAGACAACATCCGTATCAATGGTTGTGCCATTCAGTCAGAGTCACCACA 1113
          |||| |||||| || || ||| |||| |||| ||| ||| ||||| |||| |||| || ||||
Db    365 GGCCTCAAGCAAGAGACAGACAACGGCTTCCGCCCAGATATCGCCCAGATACCGGAACTATCACCGCTCACCA 424

Qy   1114 GAAGATCCAAACAACGGCTTCCGCCCAGATATGCCGGAACTATCACCGGTACCGCTCACCA 1173
          |||| ||| ||| ||||| |||| |||||| || ||| |||| ||| ||||| ||||||
Db    425 GAAGACCCCTTCAGCTGGGTTTCAGCCTGATACTGGTCGCGGATTGAGGTTTCCGGAGTGGA 484

Qy   1174 GGCGGAGCTGGGCGTCGTCTTGACGGTG---CAGCCGCAGGCTGCGGTGGCGAAATCACCGCCA 1230
          || ||||| |||| ||| ||||| |      ||||| |  || |||||||| |||||||||
Db    485 GAAGGAATGGGAATCCGACTAGACAGTGTCTTCAAGGAGGCCGTCATTTCCCCA 544

Qy   1231 CACTTTGACTCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTT 1290
          ||| ||| || |||| ||| || |||| || || || ||| ||||| |||| || ||
Db    545 CATTATGACTCCCTTCGTGGTCAAGGTTATTGCACACGGCAAAAGACCATCAGTTGCCGCC 604

Qy   1291 GCTCGTGCACAGCGCGCCGTTGGCTGGTTC 1320
          ||||| || ||||| || ||||| |||||
Db    605 ACAAAAATGAGCAGCGAGCTTTTGGGCAGAGTTC 634
```

```
RESULT 10
BF781506
LOCUS       BF781506                887 bp    mRNA    linear    EST 12-JAN-2001
DEFINITION  602104410F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4222419
            5', mRNA sequence.
ACCESSION   BF781506
VERSION     BF781506.1  GI:12086542
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 887)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM99809 row: d column: 04
            High quality sequence stop: 752.
FEATURES             Location/Qualifiers
     source          1..887
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4222419"
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      188 a    282 c    249 g    168 t
ORIGIN

Query Match         5.7%;  Score 197.4;  DB 12;  Length 887;
Best Local Similarity  56.0%;  Pred. No. 2.4e-41;
Matches 394;  Conservative   0;  Mismatches 306;  Indels   3;  Gaps   1;

Qy    746 TACACCTTTATGAACGTGACTGCTCACTGCGACCTCGTCACCAAAAAGTTGTCGAAATTG 805
              ||||||| |  |||| ||  || |||  |||||||||| ||| || ||||  |||||
Db      6 TGCACCTGTACGAGCGAGACTGCTCCATCCAGCGTCGGCACCGAAGGTGGTAGAGATCG 65

Qy    806 CGCCAGCACAGCATTTGGATCCAGAAACTGCGTGATCGCATTTGTGCCGATGCAGTAAAGT 865
              ||||| || |||||||| ||  |||  |||||||| |||||| |||||| |||| |||
Db     66 CCCCTGCTACCCACCTGGATCCCCAACTTCGCTCGCTCACGTCTCACCAGTGACTCGTCAAC 125

Qy    866 TCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCTGGCGGCTGGAATTCTTGGTCGATGAAAAAGG 925
              |||| || || || ||||| |||||||| || ||||| ||| |||||| ||||| || |||||
Db    126 TTGCCAAGCAGGTAGGCTATGAGAACGCCGGCACTGTTGGAGTTCCTGGTCGTGGACAAGCACG 185

Qy    926 GCAACCACGTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAG 985
              ||| ||| || ||||||||  ||| |||| |||  || ||||||| ||| || |||
Db    186 GCAAGCACTACTTCATCGAGGTAGAGGAGAACATCCGCATTCCCGCCTGCAGGTGAGCAGCACAGGTCACCGAGG 245

Qy    986 AAGTCACCGAGGTGGACCTGGTGAAGGCCGAGATGCGCTTGGCTGCTGGTGCAACCTTGA 1045
              |||| ||| ||||| ||||| |||| |||||| |||||| ||||||| |||||||||
Db    246 AGATCACAGATGTGACCTGGTCCATGCTCATGCTCAGATCCAGAATCCACGTGTCCGAAGGCCGAGCCTGC 305

Qy   1046 AGGAATTGGGTCTGACCCAAGATAAGATCAAGACCACCGGTCGCAGCACTGCAGTGCCGCA 1105
              ||| ||||| ||| ||| ||||| |||| |||| ||| |||| |||||| ||||||
Db    306 CTGACCTGGCCTGCGCGAGGAGAACATCCGCATCTGTGCCATTCAGTGTCGGG 365

Qy   1106 TCACCACGGAAGATCCAAACAACACGGCTTCCGCCCAGATACCGAACTATCACCGCGTACC 1165
```

```
                          |||||||| |||   ||  ||   ||||||| ||||| |||   ||  |
Db    366 TCACCACCGAGGACCCTGCACGCAGCCTTCCAGCCAGCACCAGGACACCAGGGCCGGCCATTGAGGTTTTCC 425

Qy   1166 GCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGTGCAGTCGGTGGCGAAATCA 1225
              ||    ||  ||||| |||||||| |||  | ||||  |  |  ||| || |||||
Db    426 GGAGTGGTGAGGGCATGGGCATCCGCCTGACAACGCCTCTGCATTCCAGGCGCTGTCA 485

Qy   1226 ---CCGCACACTTGACTCACCATGCTCGTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAAA 1282
                 ||||||| ||| |||||  || |||| |||| || |||||| |||| || |
Db    486 TATCGGCCCCACTATGACTCTCTCGTCGTCAAGGTCATTGACACGGCAAAGACCACCCCA 545

Qy   1283 CTGCGTGCTGCTCGTGCACAGCGGCGCGTGGCTGAGGTTCACCGTGTCGTGGTTGCAACCA 1342
              ||||  |||  ||  |||  |||| || || ||  |  || ||| ||||||| ||||||
Db    546 CAGCTGCCACCAAGATGAGCAGAGCCCTGGCCGAGTTCCGTTCCGAGGTGTAAAGACCA 605

Qy   1343 ACATTGTTTTCTTGCCGCGTTGCTGTCGGGGAAGAGGACTTCACTTCCAAGGCCATCGCCA 1402
              ||||| |||| | |  |||| |   ||| ||||||| || |||| ||||||| || || |
Db    606 ACATCCCCTTCCTGCAGAATGTTCTCAACAACACCAGCAGTTCCTGGCCAGGCACAGTGGACA 665

Qy   1403 CCGGATTATCGGCGATCACCCACACCTCCTTCAGGCTCCACC 1445
              ||   | |  |||||  |  | || || ||| ||| ||||||
Db    666 CCCAGTTCATCGATGAGAACCCTGAGCGTGTTCCAGCTTCGGCC 708

RESULT 11
AU202878
LOCUS       AU202878                515 bp    mRNA    linear    EST 17-JUL-2001
DEFINITION  AU202878 unpublished oligo-capped cDNA library, stage L2
            Caenorhabditis elegans cDNA clone yk813g01 5', mRNA sequence.
ACCESSION   AU202878
VERSION     AU202878.1  GI:14833043
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
ORGANISM    Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 515)
AUTHORS     Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
TITLE       A complementary view of the C.elegans genome
JOURNAL     Unpublished (2001)
COMMENT     Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..515
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone="yk813g01"
                     /clone_lib="unpublished oligo-capped cDNA library, stage
                     L2"
                     /sex="Hermaphrodite"
                     /tissue_type="whole animal"
                     /dev_stage="L2"
                     /note="The AD-wrmcDNA library was generated with poly(A)+
                     RNA isolated from both hermaphrodite and male N2 worms of
                     all larval stages, embryos, adults and dauers and the
                     subsequent generation of cDNAs by poly(A) priming. The
                     cDNAs were cloned into pPC86"
BASE COUNT      146 a    103 c    124 g    142 t
ORIGIN

Query Match         5.7%;  Score 196.6;  DB 9;  Length 515;
Best Local Similarity  62.9%;  Pred. No. 2.7e-41;
Matches 304;  Conservative   0;  Mismatches 179;  Indels   0;  Gaps   0;

Qy    640 GAAGCTGAAGCGGCATTCGGCGACGTTCGGTATATGTGAACGTGCTGTGATTAACCCC 699
```

inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+
enriched RNA The cDNA was ligated into Lambda gt11 from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce
E. coli Y1090 and phage DNA was purified from a liquid
lysate. The cDNA inserts were gel purified after EcoRI
digestion and ligated into pBluescript SK+. Aliquots of
the ligation were used to transform E. coli DH5alpha which
were plated onto medium with X-gal for selection of
recombinants."

BASE COUNT   188 a   312 c   241 g   208 t
ORIGIN

Query Match   5.6%;   Score 196.2; DB 14;   Length 949;
Best Local Similarity 54.7%;   Pred. No. 5.2e-41;
Matches   412; Conservative   0; Mismatches 338; Indels   3; Gaps   1;

Qy   2122 ACCCTGGATTACTACCTAAAGATGGCAGAGAGAGATCGTCAAGTCTGGCGCTCACATTCTG 2181
          | |||| ||| |||| |||||| || | | || | ||| | ||| || ||||||
Db   1 AACCTCCAGTACTACCTGGTCGACAAGCTCGTTGCCCTGCCGACATCCACGTTCTC 60

Qy   2182 GCCATTAAGGATATGGCTGGTCGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCA 2241
          ||| ||| || ||||| | | ||||| |||| | || |||| |||| || ||||
Db   61 GGCATCAAGGACATGGCCGGTGTCCTCAAGCCCACGCCGCGTACCCTTCGATCGGAGCC 120

Qy   2242 CTGGCGCGTGAATTC---GATCTGCCAGTGACGTGCACACCGACACTGCGGGTGGC 2298
          | | | |||| || |||| ||| || || | || | || |||| ||
Db   121 ATCCGCAAGGAGTACCCGACCTCCTATTCACGTTCATACTCACGATTCTCTGGTACG 180

Qy   2299 CAGTTGGCTACCTACTTGTCGCAGCTCAAGCTGCGGAGCGCCGGTGCTGCGTGCGTTCC 2358
          || || || |||| | || ||| ||||| || || | ||| ||| |||| ||||
Db   181 GGTGTTGCTTCTATGGTTGCTTCGCGGCACCACCTCCGTGGACGCCGGAGCGTGCCACC 240

Qy   2359 GCACCACTGTCTGGCACCACTCCCAGCCATCCGTCTGGCCATTCCTGTGCTGTCGCATTCGCG 2418
          | ||| |||| | |||||| | ||| || | |||| || | | |||| | ||| |||
Db   241 GACAGCTTGTCCGGCATGACCTCCCAGCCCAGCAATCCAACGCCATAATTGCTTCTCTCGAA 300

Qy   2419 CACACCCGTCGCGATACCGGTTGAGCTGCTGTTTCTGAGCCGTGTTTCTGAGCCGTACTGG 2478
          | | || | | || | || | | || | | |||| ||| ||||
Db   301 GGCATGGTGAAACGACACCGGCCTGAACCTGCGTTCGCACGTTCGCCCTGGACACCTATTGG 360

Qy   2479 GAAGCTGTGGCGGACTGTACCTGCCATTGAGTCTGGAACCCCCAGCCCAACCGGTCGC 2538
          || | || | | | | || ||||| ||| || | || ||| || ||| ||| |
Db   361 TCTCAGCTGCGTCTTCTCACTTCACCAACTTTCGAGGCTCATCTCGTCTGTCTGACCCGTAG 420

Qy   2539 GTCTACCGCCACGAAATCCAGGCGGACAGTTGTCCAACCTGGTTGTTGGCGTGCACAGGCCACCGCA 2598
          | || ||| | || || | | || || || || | | || ||| |||| | |
Db   421 GTGTACGAGCATGAGATCCCCGTGGCCAGCTTACCACATGATGTTCCACACTGCTTCCAAGCAGCTCAG 480

Qy   2599 CTGGGCCTTCTGCTTCGAGCTTCATCGAGCAGCTACGCAGCAGCCGGTTAATGAGATG 2658
          |||| ||| ||| | || | | | | | ||| || ||| || || ||
Db   481 CTTGGTCTTCTTGGAACACAATGGGCGGAGACCAAGAAAGCGTACGAGCCACGACCTC 540

Qy   2659 CTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTGGCCGACCTCCGCACCTCCAC 2718
          |||| | | ||| || | ||| | | |||| || || ||| | |||| | |||| ||
Db   541 CTTTGGCGACATTGTCAAGGTCACTCCCCACTTTCCAAGGTTCGTTGGCGACTTGGCCCAGTTC 600

Qy   2719 CTGGTTGGTGGCGGGTGTAGATCCAGAGACTTTGCTGCAGACCCACAAAAGTACGACATC 2778
          ||||| || || | | || ||| | | | | | | | |||| | | | | |||
Db   601 ATGGTTTCGAACAAGCTTACCCCTGGAGGATGTCAAGGCGCGCGCGGAGCTCGACTTC 660

Qy   2779 CCAGACTCTGTCATCGGCGGTTCCTGCGGCGGGAGCTTGGTAACCTCCAGGTGCTGGCCA 2838
          |||| ||| ||| | | || ||| ||||| || | ||||| | ||| || ||||
Db   661 CCCGGACATTGTCAAGGTCACTCCCGAGTTCCTTGAGGGTATGATGGGCCAGCCATTTGGTGGATTCCCC 720

Qy   2839 GAACCACTGCGCACCCGCCGCACTGGAAGGCCGC 2871
          || || | | | | || || | ||| | | ||
Db   721 GAGCCCCTTCGCTCCAAGCGTCTTCCCGTGGCCGC 753

Db   32 GAAGCTCAAGCTGCGGTTTGAGAGGACGGAAGTCTTTTCGTTGAGAAAGTTGTTGAGAGACCA 91

Qy   700 CAGCACATTGAAGTGGAGATCCTTGGCGATCGCACTGCACCCAAAAAGTTGTACAACCTTTATGAA 759
          |||||| || |||| || || |||| | || ||| || ||| ||| || |||| |||||
Db   92 AGACATATTGAAGTTCAGCTGCTGCTTGGAGACCATCATGGAAATATATTGTTCATTGTATGAG 151

Qy   760 CGTGACTGCTGCACTGCAGCGGTCGTCACCAAAAAGTGTCGAAAATTGCGCCACCACAGCAT 819
          || || || | | ||||| || ||| | ||||| | ||| |||| | || || || || |
Db   152 CGTGATTGTTCAGTGCAACGTCGTCATCATCAAAAGGTGTTGAAATTGCTTCCAGGCCAGCT 211

Qy   820 TTGGATCCAGAACTGCCTGATCGATCGCATTTGTGCGGATGCGGATCGCAGTAAGTTCTCGCCGCTCCATT 879
          |||| || | || || | |||| | ||| || || | | || | || | ||| ||| |
Db   212 CTCCCAGAGGTGTTCGTGAGAAAATTTTGGCAGACGCTCTTCGACCTTGCAGACGCTCGCAACGCATGTT 271

Qy   880 GGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGTCGATGAAAAGGGCAACCACGTTTTC 939
          || |||| || | | | ||| | ||| | | | || || | | || || |
Db   272 GGATACCAAAATGCTGGTACGCGTCCTCAAGTTCCTGGTTGATCAGAGGGCCAACTACTATTTC 331

Qy   940 ATCGAAATGAACCCACGTATCCAGTTGAGCACCACGTGACTGAGAAGTCACCGAGGTG 999
          || || | | || |||| ||| | | | || | | || || | |||| |
Db   332 ATCGAAGTGAATGCACGTCTTCAAGTCGAGCATACAGTAACTGAGAGATCACTGGTGTG 391

Qy   1000 GACCTGGTGAAGGCCGAGATGCGCTTGGCTGGTCGCTGTGCGTGGCAAGGCGGTGCTG 1059
          ||| |||| || | | | | |||| | | || || || || | ||| ||
Db   392 GATCTTGTCCAAGTCAAATTCGTATCGGCCGAAGGAAAATTCTCTGGATGATCTGAAGCTT 451

Qy   1060 ACCCAAGATAAGACTCAAGACCCACCGTGCAGCACTGCAGTGCCGCATCACCACGGAAGAT 1119
          || ||| || | || | || || | ||||| || | | || | | ||| |||
Db   452 TCACAGGAAACTATTCAAACTACTGCTGCTATTCAATGTCGTGTCAACTGAAGAT 511

Qy   1120 CCA 1122
          |||
Db   512 CCA 514

RESULT 12
BQ752438

LOCUS   BQ752438   949 bp   mRNA   linear   EST 18-JUL-2002
DEFINITION EST633001 DSCT Colletotrichum trifolii cDNA clone pDSCT12-67, mRNA
            sequence.
ACCESSION   BQ752438
VERSION   BQ752438.1 GI:21907843
KEYWORDS   EST.
SOURCE   Colletotrichum trifolii.
  ORGANISM Colletotrichum trifolii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
            mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE   1 (bases 1 to 949)
  AUTHORS   Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
            Cheung,F. and Fraser,C.M.
  TITLE   ESTs from mycelia of Colletotrichum trifolii race 1
  JOURNAL   Unpublished (2002)
COMMENT   Other ESTs: EST633000
            Contact: Deborah A. Samac
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
            Tel: 612 625 1243
            Fax: 651 649 5058
            Email: debbys@puccini.crl.umn.edu
            TIGR sequence name: MTSAL67TV More information is available at:
            www.medicago.org
            Seq primer: (gtA AtA CgA CtC ACt AtA ggg C).
FEATURES   Location/Qualifiers
     source   1..949
                     /organism="Colletotrichum trifolii"
                     /strain="race 1"
                     /db_xref="taxon:5466"
                     /clone="pDSCT12-67"
                     /clone_lib="DSCT"
                     /tissue_type="mycelia"
                     /dev_stage="Young, actively growing mycelia (3 days after

RESULT 13
BI605019
LOCUS   BI605019   666 bp   mRNA   linear EST 07-SEP-2001
DEFINITION RH70532.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH70532 5 similar to
            BcDNA:GH06348: FBan0001516 GO:[pyruvate carboxylase (GO:0004736)]
            located on: 2R 46B14-46C1;, 08/24/2001, mRNA sequence.
ACCESSION   BI605019
VERSION   BI605019.1 GI:15500544
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 666)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Paoleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE   BDGP/HHMI RH Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   BDGP
            ,Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003831: arm:2R [4876906,5137815]
            estimated-cyto:46B13-46E4: 08/24/2001
            Plate: RH.705 row: C column: 8
            High quality sequence stop: 521.
FEATURES   Location/Qualifiers
source   1..666
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
            /clone="RH70532"
            /clone_lib="RH Drosophila melanogaster normalized Head
            pFlc-1"
            /sex="male and female"
            /dev_stage="Adult"
            /lab_host="DH5-alpha TonA"
            /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
            BamHI; Library was kindly generated by Piero Carninci at
            the RIKEN. The library was normalized and excised using
            Cre recombinase. Plasmid cDNA library."
BASE COUNT   142 a   195 c   207 g   121 t   1 others
ORIGIN

Query Match   5.6%; Score 195; DB 13; Length 666;
Best Local Similarity 58.2%; Pred. No. 8.7e-41;
Matches   361; Conservative   0; Mismatches 256; Indels   3; Gaps   1;

Qy   1561 AAGGATACTGCCACTGCCACGCGGTTCCCGTGACCGCTGAGCAGCTTGGCCAGCCGCCG 1620
          ||||||||| ||||||||| ||||||||||||||||||||||||||| ||||||
Db   46 AAAGTTACGGAACCACCGAAGGACTCCGCGAGGTCCTGGTGTGCGAAGGTCCGGGAGCC 105

Qy   1621 TTTGCTGCTGTGATCTCCGTGAGCAGCAGGACCACTGGCAGTTACTGATACCACCTTCCGGCAT 1680
          ||||| |||||||||| ||| ||||||||||| |||||||||||||| |||||||||| ||
Db   106 TTCGGCCAAGGAGGTGCGCAACCGTAAGGAACTGCTACTCATCGAACACCACCCTTCCGGCAT 165

Qy   1681 GCACACCCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
          ||||||||| |||||| |||||||||| ||||||||||| |||||||||||||||||
Db   166 GCCCACACGTCGCTCTGCGGCCACCCGCGTGCGTTCCCACGATCTGCTGAAGATCTCCCCC 225

Qy   1741 GCCGTCGCAAAGCTGACTCCGTGAGCAGGAGCAGGAAGCCTGGGCGGCGGCGCGACCTAC 1800
          |||||| ||||||||||| |||||||||||||||||||| |||||||||| ||||||||||
Db   226 TACGTGACGCACAAGTTCAACAACCTGTATTCGCTGGAGAACTGCGGAGAACTGGGCGGGCGGGAGCGCCACCTTC 285

Qy   1801 GATGTGGCGATGCGGTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCGTGCGCGAG 1860
          |||||||||||| ||| ||||| |||||||||||
Db   286 GACGTGGCGGCTGCCGCTTCCTGCGCCGATGACGAGTGCCTGACGGCGAGATGGGCCAAG 345

Qy   1861 GGCATGCCGAATGTAAACATTCAGATGCTGCTTCGCCGGCCGCCAACACCGTGGATCACCTT 1920
          ||||||||| ||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   346 CGCATCCCGACATTCCCTTCCCAGATGCTGTTGCGCCGGAGCAGCCGCCGGCGCCGGCCGCTACACC 405

Qy   1921 CCGTACCCAGACTCCGTCTGCCGCGCGGTTTGTTAAGGAAGCTGCCAGCTCCGGCTCGGAC 1980
          |||| ||||||| |||| |||| ||||||||| |||||||||||||||||||||||||||||
Db   406 AGCTATCCGGACAACGTGGTCTACAAGTTCTGTGGAGCGTGGTCGTGCGAGACCGGCATGGAC 465

Qy   1981 ATCTTCCGCATCTTCGACGCGGCCTTAACGACGTCTCCCAGATGGGTCCAGCAATCGACGCA 2040
          |||||||| || ||||||||||||| |||||||||||||||||
Db   466 ATCTTCAGGGGTGTTCGACTCGCTCAACTACCTGCCCAACCTGATCCTCCGGCCATGGAAGCC 525

Qy   2041 GTCCTGGAGACCAACACCGCGTAGCCGAGGTGGCTTATTCTGGTGATCTTCTCT 2100
          |||||||||||||| |||||||||||||||| |||||||| |||||| ||||||
Db   526 TGCTGAAAGGCGGGCGGCCGCGTGGT---GGAGGCTGCCATCTTCTATACCGGAGACGTCAGC 582

Qy   2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTC 2160
          ||||| |||| ||||||||| ||| |||||||||||| ||||| ||||||||| ||||||||
Db   583 GATCCCCAGCGCACCAAGTATGATCTGAAAATACTACACTAACCTTGCCGATGGAGCTGGTC 642

Qy   2161 AAGTCTCGGCGCGCTCACATTCT 2180
          ||| ||||| |||| ||||
Db   643 AAGGCCGGGCACCCACGTGCT 662

RESULT 14
CNS06M6L
LOCUS   CNS06M6L   1030 bp   DNA   linear GSS 17-JUN-2001
DEFINITION T3 end of clone AU0AA002D08 of library AU0AA from strain CBS 3082
            of Saccharomyces kluyveri, genomic survey sequence.
ACCESSION   AL405091
VERSION   AL405091.1 GI:12167396
KEYWORDS   GSS.
SOURCE   Saccharomyces kluyveri
ORGANISM   Saccharomyces kluyveri
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1 (bases 1 to 1030)
AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuvglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weisenbach,J.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE   20584711
PUBMED   11152876
REFERENCE   2 (bases 1 to 1030)
AUTHORS   Neuvglise,C., Bon,E., Lepingle,A., Wincker,P., Artiguenave,F.,
            Gaillardin,C. and Casaregola,S.
TITLE   Genomic exploration of the hemiascomycetous yeasts: 9.
            Saccharomyces kluyveri
JOURNAL   FEBS Lett. 487 (1), 56-60 (2000)
MEDLINE   20584719
PUBMED   11152884
REFERENCE   3 (bases 1 to 1030)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

```
FEATURES           5 kb were prepared and both extremities were sequenced. See
                   keywords for description of this sequence and for the sequence of
                   the other extremity of this insert.
        source     1..1030
                   /organism="Saccharomyces kluyveri"
                   /strain="CBS 3082"
                   /db_xref="taxon:4934"
                   /clone="AU0AA002D08"
                   /clone_lib="AU0AA"
                   /note: "end : T3"
        misc_feature  <12..>1029
                   /note="similar to Saccharomyces cerevisiae ORF YBR218c [
                   PYC2 ; pyruvate carboxylase 2 ]
                   2 putative frameshift(s)
                   similar to Saccharomyces cerevisiae ORF YGL062w [ PYC1 ;
                   pyruvate carboxylase 1 ]"
                   /evidence=not_experimental
BASE COUNT     270 a    247 c    230 g    282 t     1 others
ORIGIN

Query Match          5.6%;   Score 193.2;   DB 17;   Length 1030;
Best Local Similarity  53.3%;  Pred. No. 3.5e-40;
Matches  523;   Conservative     0;   Mismatches  423;   Indels   36;   Gaps    4;

Qy    892 GCGGGAACCGTGGAATTCTTGGTCCGATGAAAAGGCAACCACGTTTCATCGAAATGAAC  951
              ||| ||||||  |||||| || |||||| || ||||| ||| |||||||| |||||||
Db     21 GCCGGTACTGCCGAGTTTTTTGGTCGATGCGCAAAACAGACATTATTTCATCGAAATCAAC   80

Qy    952 CCACGTATCCAGGTTGAGCACACCGTGACTGAAGAGTCACCGAGGTGGACCTGGTGAAG 1011
              ||||||  |||||||||| ||||||| || |||||  |||||  || |||| |||||||
Db     81 CCAAGAATCCAAGTCGAACATCATCACCATCACTGAAGAAATCACCGGTCGTCGACGTTCGGGCT  140

Qy   1012 GGCGCAGATGCGCTTGGCTGCTGGTGCAACCTGAAGGAATTGGGTCTGACCCCAAGATAAG 1071
              ||||||||||| ||||| ||||||| |||||||||| ||||| || ||||| ||||||
Db    141 GCTCAAATTCAAATTGCCGCGCAGCCGTTCCTTGCAACAACTAGGCTTGTTGTTGCAGGATAAA  200

Qy   1072 ATCAAGACCCACGGTGCAGCACTGCGCGAGTGCCGATCACCACGGAAGATCCAAAACAACGGC 1131
              ||||| || || ||| ||  || ||||   |||  |||||||| |||| || ||||||  ||
Db    201 ATAACCACTCGTGGGTTTGCCATTCAATGTCGTATAACCACTGAAGAATCCTTCAAAGAAC  260

Qy   1132 TTCCGCCCAGATACCGGAACTATCACCGGTACCGGTCACCAGGCGGAGCTGGCCTTCGT 1191
              ||||| || ||||| |||||| |||| ||||||||| || |||| |||||||  |||
Db    261 TTCCAACCTGATACCGGTCGTATCGAAGTCTATCGTCCCGTGGTTGGTAATGGTGTCAGA  320

Qy   1192 CTTGACGGGTG---CAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTG 1248
              ||||| || ||   ||||| |||| ||||| ||||| ||| ||||||| || ||||||
Db    321 CTAGATGGTTGGTAACGCTTACGCAGGTGTCTGCATTTCTCCACATTACGACTCTATGTTA  380

Qy   1249 GTGAAAATGACCTGCCGTGGTTCCGACTTTGAAACTGCTGTTGCTCGTGCACAGCGCGCG 1308
              |||| ||||||| || || ||| ||||||||||   |||  | || || ||||| ||| ||
Db    381 GTCAAGTGTTCCTGTTCCGGCTCAACCTCAACCTACGAAATTGTTCGTCGTGAAGATGTTGCGTGCC  440

Qy   1309 TTGGCTGAGTTCACCGTGCTGTGGTGTTGCAACCACATTGGTTTCTTGCTGGCGTTGCTG 1368
              |||||    ||| ||||| ||| ||  || || |||  ||||| |||| || |||| ||||
Db    441 TTGATAGAATTCAGAATCAGAGGTGTTAAGACCACTATTCCTTTCCATTCCTTTCTTCTACTTTGCTA  500

Qy   1369 CGGGAAGAGGACTTCACTTCCAAGCGCATCGCACCGGATTTATCGGCGATCACCCCACAC 1428
              || ||||| ||  || ||| |||   ||| ||||||| ||| ||  |||| |||||||||
Db    501 ACCCATTCCGTGTTCAAGACGGTGAATATTGGACTACTTTCATCGACGACACTCCACAA  560

Qy   1429 CTCCTTCAGGCTCCACCTGCGGATGAGATGAGAGGAGACGCATCCTGGATTACTTGGCAGAT 1488
              ||||| ||| ||   || ||| || ||||| |||| || ||||| |||||| |||||
Db    561 CTATTTGAAATGGTTTCTTCTCAAAACAGAGCCCAAAAGTTGCTACACTACCTGCCGAC  620

Qy   1489 GTCACCGTGAACAAGCCTCATGGTGTCGGTCAAAGGATGTTGCAGCACCAATC-----G 1543
              || ||||| ||  || ||  || |||  || ||| |||  |||||||||| ||
Db    621 TTGGCTGTCAATGGTTCTTCGATCAAGGGTCAAATTGGCCTACCAAAATTGACCACTCAG  680

Qy   1544 ATAAGCTGCCCAACATCAAGGAATCGCCACTGCCACGCGGTTCCCGTGACCGGC------ 1596
              |||| |||  ||| ||||| ||   |||||| |  || ||| || ||| ||||
Db    681 CCAAGCGTTCCAACTTTACACGACTCAAACGGTAACGTTGTCGACGTTGTTGGCTGTGCGAC  740
```

```
RESULT 15
AI293242
LOCUS           AI293242                   625 bp    mRNA    linear    EST  19-APR-2001
DEFINITION      GH16378.5prime GH Drosophila melanogaster head pOT2 Drosophila
                melanogaster cDNA clone GH16378 5prime, mRNA sequence.
ACCESSION       AI293242
VERSION         AI293242.1  GI:3942649
KEYWORDS        EST.
SOURCE          fruit fly.
  ORGANISM      Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE       1  (bases 1 to 625)
  AUTHORS       Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
                Lewis,S. and Rubin,G.M.
  TITLE         BDGP/HHMI Drosophila EST Project
  JOURNAL       Unpublished (2001)
COMMENT         Contact: Stapleton, M.
                BDGP
                Lawrence Berkeley National Lab
                One Cyclotron Rd, Berkeley, CA 94720, USA
                Fax: 510 486 6798
                Email: http://www.fruitfly.org/EST., est@fruitfly.berkeley.edu
                Plate: 163  row: G  column: 6
                High quality sequence stop: 473.
FEATURES             Location/Qualifiers
        source       1..625
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GH16378"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2. Plasmid cDNA library."
BASE COUNT     134 a    196 c    176 g    119 t
ORIGIN

Query Match          5.5%;   Score 192.6;   DB 9;   Length 625;
Best Local Similarity  60.2%;  Pred. No. 3.7e-40;
Matches  356;   Conservative     0;   Mismatches  229;   Indels    6;   Gaps    2;

Qy   1763 AGCTTTGTCCGTGGAGGCCTGGGCGCGGCCGCGACCTACGATGTGGCGATGCGTTCCTCT 1822
              || ||||  |||| ||| ||||| || || || || ||||  || || |||| |||||
Db     37 ACCTGTATTCGTCGGAGAACTGGGCGGGCGGCCACCTTCGACGTCGACGTGGCGTGCTTCTGC   96

Qy   1823 TTGAGGATCCGTGGGACAGGCTCGACGAGCTCGCCGAGGCGGCGATGCCGAGGCTAAACAATTC 1882
              ||  |||  || ||| || || ||| || ||||| ||  ||| ||| || ||||||
Db     97 ACGAGTGCCCCGTGGGAGCCCTGGAGGAGGATGCGCAAGCGCCATCCCGAACATTCCCTTCC  156
```

```
Qy   1883  AGATGCTGCTTCGCGGCGCCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCC  1942
                 ||||||||   |||| ||||||  |||||| || ||| |||| |||||||  |||
Db    157  AGATGCTGTTGCGCGGGAGCCAACGCCGTCGGCTACACCAGCTATCCGGACAACGTGGTCT  216

Qy   1943  GCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCTGGACATCTTCCGCATCTTCGACCGCGC  2002
                  |||  |  ||||||||||||| ||||  |  |||||||| || |||||| |||||||  ||
Db    217  ACAAGTTCTGTGAGCTGGCTGTGCAGACCGGCATGGACATCTTCAGGGTGTTCGACTTCGC  276

Qy   2003  TTAACGACGTCTCCCAGATGGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGG  2062
                 | || |||||| |||||||| | || ||| |||| ||| ||    |      || ||||
Db    277  TCAACTACCTGCCCAACCTGATCCTCGGCCATGGAAGCCGCTGGAAAGGCGGGCGCGTGG  336

Qy   2063  TAGCCGAGGTGGCTATGGCTTATTCTGGTGATCCTCTGATCCAAATGAAAAGCTCTACA  2122
                  |||| ||| |||| || |||| ||| |||||||| |   ||||| |||||||||| ||
Db    337  T---GGAGGCTGCCATCTCCTATCCGGAGACGTCAGCGATCCCAAGCGCACCAAGTATG  393

Qy   2123  CCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTGGCGCTCACATTCTGG  2182
                  |||||| |  | || || |   ||  |||  |||| | |||  ||  ||| |||   |||
Db    394  ATCTGAAATTACTACACTAACCTTGCCGATGAGCTGGTCAAGGCGGCACCCACGTGCTCT  453

Qy   2183  CCATTAAGGATATGGCTGGCTCGTCTTCGCCCAGCTGCGTAACCAAGCTGGTCACCGCAC  2242
                  |||||||||||||||||  | |     |  |||| | || |   ||||  ||||  ||
Db    454  GCATAAAGGATATGCTGGTGGTCTGTGCTCAAGCCCGAATCCGGCTTCTGATCACAGCCA  513

Qy   2243  TGCG--CCGTGAATTCGATCTGCCAGTGCACGTGCACACCCACGACACTGCCGGGTGCC  2299
                 ||||     |  |  |||||| |||||  ||  ||  |||||||||||  ||||||  ||
Db    514  TTCGTGCAAGCACCCCGACATCCCAATCCATATCCACACCCACGACACCCTCTGGAGCGG  573

Qy   2300  AGTTGGCTACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTTGTTGACG  2350
                 ||||     ||   |  ||||| |||| |  | |||| ||  ||||| ||||
Db    574  GAGTGGCCTTCATGCTGGCATGCGCCAATGCCGGAGCCCGATGTGGTGGACG  624
```

Search completed: April 5, 2003, 07:14:38
Job time : 5257 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April   5, 2003, 05:32:17 ; Search time 126 seconds
                                     (without alignments)
                                     8455.516 Million cell updates/sec

Title: US-09-974-973-1
Perfect score: 3474
Sequence: 1 gtgactgctatcaccctgg.........tgatcgtcgtcgttcctaa 3474

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
         1: /cgn2_6/ptcdata/1/ina/5A_COMB.seq:*
         2: /cgn2_6/ptcdata/1/ina/5B_COMB.seq:*
         3: /cgn2_6/ptcdata/1/ina/6A_COMB.seq:*
         4: /cgn2_6/ptcdata/1/ina/6B_COMB.seq:*
         5: /cgn2_6/ptcdata/1/ina/PCTUS_COMB.seq:*
         6: /cgn2_6/ptcdata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	3398.8	97.8	3621	4	US-09-220-081-1	Sequence 1, Appli
2	3398.8	97.8	3621	4	US-09-677-575-1	Sequence 1, Appli
3	476.6	13.7	3465	4	US-09-134-001C-591	Sequence 591, App
4	306.6	8.8	1362	1	US-07-956-700B-5	Sequence 5, Appli
5	306.6	8.8	1362	1	US-08-476-537-5	Sequence 5, Appli
6	306.6	8.8	1362	1	US-08-485-607-5	Sequence 5, Appli
7	306.6	8.8	1362	1	US-08-475-879-5	Sequence 5, Appli
8	306.6	8.8	1362	4	US-09-433-043B-5	Sequence 5, Appli
9	305	8.8	1362	1	US-08-611-107-7	Sequence 7, Appli
10	305	8.8	1362	2	US-08-422-560A-7	Sequence 7, Appli
11	305	8.8	1362	4	US-08-468-793-7	Sequence 7, Appli
12	302.2	8.7	3077	1	US-08-074-121-1	Sequence 1, Appli
13	302.2	8.7	3077	5	PCT-US94-06447-1	Sequence 1, Appli
14	297	8.5	3065	1	US-07-956-700B-1	Sequence 1, Appli
15	297	8.5	3065	1	US-08-476-537-1	Sequence 1, Appli
16	297	8.5	3065	1	US-08-485-607-1	Sequence 1, Appli
17	297	8.5	3065	1	US-08-611-107-5	Sequence 5, Appli
18	297	8.5	3065	2	US-08-422-560A-5	Sequence 5, Appli
19	297	8.5	3065	1	US-08-475-879-1	Sequence 1, Appli
20	297	8.5	3065	4	US-08-468-793-5	Sequence 5, Appli
21	297	8.5	3065	4	US-09-433-043B-1	Sequence 1, Appli
22	267.4	7.7	1921	1	US-08-662-344-1	Sequence 1, Appli
23	255	7.3	3231	1	US-08-074-121-4	Sequence 4, Appli
24	255	7.3	3231	5	PCT-US94-06447-4	Sequence 4, Appli
25	254.8	7.3	23673	4	US-09-773-816-1	Sequence 7, Appli
c 26	247.4	7.1	19702	4	US-08-961-527-7	Sequence 7, Appli
c 27	209	6.0	1236	4	US-09-134-001C-384	Sequence 384, App

28	167.8	4.8	1386	4	US-09-134-001C-767	Sequence 767, App
29	136.8	3.9	1916	3	US-08-714-916-96	Sequence 96, Appl
30	136.8	3.9	1916	4	US-09-265-315-96	Sequence 96, Appl
31	136.8	3.9	1916	4	US-09-265-315-96	Sequence 96, Appl
32	136.8	3.1	1916	4	US-09-266-417-96	Sequence 96, Appli
c 33	108.2	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
34	99	2.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c 35	88.2	2.5	762	4	US-08-998-416-387	Sequence 387, App
36	85.8	2.5	6270	1	US-08-418-893D-25	Sequence 25, Appl
37	85.8	2.4	6790	1	US-08-418-893D-22	Sequence 22, Appl
38	85	2.0	523	4	US-08-998-416-27	Sequence 27, Appl
39	68	2.0	7470	3	US-08-417-089-5	Sequence 5, Appli
40	68	2.0	7470	3	US-08-695-651-5	Sequence 5, Appli
41	68	2.0	7470	4	US-08-930-285-5	Sequence 5, Appli
42	68	2.0	7470	4	US-08-695-421-5	Sequence 5, Appli
43	68	1.7	7470	4	US-08-697-826A-9	Sequence 9, Appli
44	57.6	1.3	6993	4	US-08-934-386-6	Sequence 6, Appli
45	46.2		600	4	US-09-364-083-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-220-081-1
; Sequence 1, Application US/0922081
; Patent No. 6171833
; GENERAL INFORMATION:
;   APPLICANT: Sinskey, Anthony J.
;   APPLICANT: Lessard, Philip A.
;   APPLICANT: Willis, Laura B.
;   APPLICANT: Stephanopoulos, Gregory
;   TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
;   FILE REFERENCE: 1533.0790000
;   CURRENT APPLICATION NUMBER: US/09/220,081
;   CURRENT FILING DATE: 1998-12-23
;   NUMBER OF SEQ ID NOS: 36
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (199)..(3621)
US-09-220-081-1

Query Match   97.8%;   Score 3398.8;   DB 4;   Length 3621;
Best Local Similarity 98.6%;   Pred. No. 0;
Matches 3427;   Conservative   0;   Mismatches   47;   Indels   0;   Gaps   0;

Qy   1 GTGACTGCTATCACCCTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   148 GTGACTGCTATCACCCTGGCGGTCTCTTGTTGAAAGGAATAATTACTCTAGTGTCGACT 207

Qy   61 CACACATCTTCAACGCTTCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 CACACATCTTCAACGCTTCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC 267

Qy   121 GCGGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACCGGTGCTATTTACCCC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 GCGGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACCGGTGCTATTTACCCC 327

Qy   181 CGTGAAGATCGGGATCATTCCACCGGTCTTCGTTCGTTCGAAGCTGTCCGCATTGGTGTACT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CGTGAAGATCGGGATCATTCCACCGGTCTTCGTTCGTTCGAAGCTGTCCGCATTGGTGTACC 387

Qy   241 GAAGGCTCACCAGTCAAGGGCTACCTGACATCGATGAAATTATCGGTGCAGCTAAAAAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GAAGGCTCACCAGTCAAGGGCTACCTGACATCGATGAAATTATCGGTGCAGCTAAAAAA 447

Qy   301 GTTAAAGCAGATGCTATTTACCCGGGATATGGCTTCCTGTCTGTCGAAAATGCCCAGCTTGCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

This page contains a dense two-column DNA sequence alignment between a query (Qy) and database (Db) sequence, with matching positions indicated by vertical bars. The alignment blocks are labelled with "Qy" and "Db" and numbered by sequence position.

Right column (positions ~1441–2667):

```
Qy 1441 ...TGATGAGCAGGAGGACGCGATCCTGGATGTCACCGTGAAC 1500
Db 1588 ...TGATGAGCAGGAGGACGCCATCCTGGATGTCACCGTGAAC 1647

Qy 1501 AAGCCTCATGGTGCGTCCAAAGGATGTTGCAGCAACCAATCGATAAGCTGCCAACATC 1560
Db 1648 AAGCCTCATGGTGCGTCCAAAGGATGTTGCAGCTCCTATCGATAAGCTGCCTAACATC 1707

Qy 1561 AAGGATCTGCCACTGCCACCGCCGGTTCCCGTGACCGGCTGAGCAGCTTGGCCCAGCCGCG 1620
Db 1708 AAGGATCTGCCACTGCCACCGCCGGTTCCCGTGACCGGCTGGGCCGCGCTTGGCCCAGCCGCG 1767

Qy 1621 TTTGCTCTGTGATCTCCGTGAGCAGGACGCACTGGCCAGTTACTGATACCACCTTCCGCGAT 1680
Db 1768 TTTGCTCTGTGATCTCCGTGAGCAGGACGCACTGGCCAGTTACTGATACCACCTTCCGCGAT 1827

Qy 1681 GCACACCAGTCTTTGCTTGCCACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1740
Db 1828 GCACACCAGTCTTTGCTTGCCACCCGAGTCCGCTCATTCGCACTGAAGCCTGCGGCAGAG 1887

Qy 1741 GCCGTCGAAAGCTGACTCCTGAGCTTTGTCCGTGAGGCCTGGGCGCGGCGCGCGACCTAC 1800
Db 1888 GCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGAGGCCTGGGCGCGGCGCGCGACCTAC 1947

Qy 1801 GATGTGGCGATGCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCCGCGAG 1860
Db 1948 GATGTGGCGATGCGTTTCCTCTTTGACGATCCGTGGGACAGGCTCGACGAGCTGCCGCGAG 2007

Qy 1861 GGGATGCGAATGTAAACATTCAGATGCGTGCTTCGCGGGCCGCAACACCGTGGGATACACCC 1920
Db 2008 GCGATGCGAATGTAAACATTCAGATGCGTGCTTCGCGGCCGCAACACCGTGGGATACACCC 2067

Qy 1921 CCGTACCCAGACTCCGTGCGCGCGGGTTGTTGAAGGAAGCTGCCAGCTCCGGCGTGGAC 1980
Db 2068 CCGTACCCAGACTCCGTGCGCGCGGGTTGTTGAAGGAAGCTGCCAGCTCCGGCGTGGAC 2127

Qy 1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCAGCAATCGACGCA 2040
Db 2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCAGCAATCGACGCA 2187

Qy 2041 GTCCTGGAGACAACACCGTAGCCGAGGTGGCTATGGCTTATTCTGGTATCTCTCT 2100
Db 2188 GTCCTGGAGACAACACCGTAGCCGAGGTGGCTATGGCTTATTCTGGTATCTCTCT 2247

Qy 2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGAGAGATCGTC 2160
Db 2248 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGAGAGATCGTC 2307

Qy 2161 AAGTCTGGGCGCTCACATTCTGGCCATTCTGGCATTAAGGATATGGCTGGTCTCGCTTCGCCCCAGCTGCG 2220
Db 2308 AAGTCTGGGCGCTCACATTCTGGCCATTCTGGCATTAAGGATATGGCTGGTCTCGCTTCGCCCCAGCTGCG 2367

Qy 2221 GTAACCAAGCTGGTCACGGGAGTTCGATCTGCCACGTGCACGGT 2280
Db 2368 GTAACCAAGCTGGTCACGGGAGTTCGATCTGCCACGTGCACACC 2427

Qy 2281 CACGACACTGCGCGGTGGCCAGTTGGCTACCTACTTTGCTCGCAAGTCTGGTGCAGAT 2340
Db 2428 CACGACACTGCGCGGTGGCCAGTTGGCCAACCTACTTTGCTCGCAAGTCTGGTGCAGAT 2487

Qy 2341 GCTGTTGACGGTGCTTCCGCACCTGTCTCGCCACCTCCCAGCCATCCCGTCTGCC 2400
Db 2488 GCTGTTGACGGTGCTTCCGCACCACCTGTCTCGCCACCTCCCAGCCATCCCGTCTGCC 2547

Qy 2401 ATTGTTGCCATTCGGCCACCGGATAACCGGTTTGAGGCTTCGAGGCTGTTTCT 2460
Db 2548 ATTGTTGCCATTCGGCCACCGGATAACCGGTTTGAGGCTTCGAGGCTGTTTCT 2607

Qy 2461 GACCCTGCCGAGCCTGACCTGGGAAGTCTGGCGCGCACTGTACCTGCCCATTTGAGTCTGAACC 2520
Db 2608 GACCCTGCCGAGCCGGTACTGGGAAGCAGTGCGCGGACAGTGTACCTGCCCATTTGAGTCTGGAACC 2667
```

Left column (positions ~361–1587):

```
Db  448 GTTAAAGCAGATGCCATTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCC 507
Qy  361 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGCCCAACCCCAGAGGTTCTTGATCTC 420
Db  508 CGCGAGTGCGCGGAAAACGGCATTACTTTTATTGCCCAACCCCAGAGGTTCTTGATCTC 567

Qy  421 ACCGGTGATAAGTCTCGTGCGTAACCGCGCGAAGAAGGCTGGTCTCGCCAGTTTTGGGG 480
Db  568 ACCGGTGATAAGTCTCCGCGCGGTACCGCGCGAAGAAGGCTGGTCTCGCCAGTTTTGGCG 627

Qy  481 GAATCCACCCGAGCAAAAACATCGATGACATCGTTAAAAGCGTGAAGGCCAGACTTAC 540
Db  628 GAATCCACCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGTGAAGGCCAGACTTAC 687

Qy  541 CCCATCTTTGTAAAGGCAGTTGCCCGTGGTGGCGGACGCGGGTATGCGCTTTGTTCTTCA 600
Db  688 CCCATCTTTGTGAAGGCAGTTGCCCGTGGTGGCGGACGCGGGTATGCGTGTTTTGTTGCTTCA 747

Qy  601 CCTGATGAGCTCCCAAAATTGGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCCGC 660
Db  748 CCTGATGAGCTTCGCAAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCTTTCGC 807

Qy  661 GACGGTTCGGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC 720
Db  808 GATGGCGCGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC 867

Qy  721 CTTGGGATCGGCACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 780
Db  868 CTTGGCGATCACACTGGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT 927

Qy  781 CGTCACCCAAAAAGTGTCGAAATTGCCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 840
Db  928 CGTCACCCAAAAAGTGTCGAAATTGCCGCCAGCACAGCATTTGGATCCAGAACTGCGTGAT 987

Qy  841 CGCATTTGTTGCGGATGCAGTAAAAGTTCGCCGCCTCCATTGGTTACCAGGGCGCGGGAACC 900
Db  988 CGCATTTGTTGCGGATGCAGTAAAGTTCGCCGCCTCCATTGGTTACCAGGGCGCGGGAACC 1047

Qy  901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTTTCATCGAAATGAACCCACGTATC 960
Db 1048 GTGGAATTCTTGGTCGATGAAAAGGGCCAACCACGTTCTTCATCGAAATGAACCCACGTATC 1107

Qy  961 CAGGTTGAGCACACCGTGACTGACGGTCACCGAGCTGGACCTGGACTGGAAGGCGCAGATG 1020
Db 1108 CAGGTTGAGCACACCGTGACTGACGGTCACCGAGCTGGACCTGGACTGGAAGGCGCAGATG 1167

Qy 1021 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1080
Db 1168 CGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC 1227

Qy 1081 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1140
Db 1228 CACGGTGCAGCACTGCAGTGCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCA 1287

Qy 1141 GATACCGGAACTATCACCCGGTACCCGGTGACTGACGGCGGCTGTCGTCTTGACCGGT 1200
Db 1288 GATACCGGAACTATCACCCGGTACCCGGTGACTGACGGCGGCTGTCGTCTTGACCGGT 1347

Qy 1201 GCAAGTCAGCTGCGGTGGCGAAATCACCCGCACTTTGACTCCATGCTCGTGGAAAATGACC 1260
Db 1348 GCAAGTCAGCTGCGGTGGCGAAATCACCCGCACTTTGACTCCATGCTCGTGGAAAATGACC 1407

Qy 1261 TGCCGGTGGTTCCGACTTTGAAACTGCTGTTTGCTCGTCGTGTCGACAGCGCCGGTGAGGTTC 1320
Db 1408 TGCCGGTGGTTCCGACTTTGAAACTGCTGTTTGCTCGTCGTGTCGACAGCGCCGGTGAGGTTC 1467

Qy 1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGGCGGGAAGAGGAC 1380
Db 1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGGCGGGAAGAGGAC 1527

Qy 1381 TTCACTTCCAAGCGCCATCGGCCCGGATTTATCGGCGACATTCACCCCACACCTTCCTTCAGGCT 1440
Db 1528 TTCACTTCCAAGCGCCATCGGCCCGGATTCATTGCCGCGATCACACCCGCACCCTTCCTTCAGGCT 1587
```

```
;  APPLICANT: Stephanopoulos, Gregory
;  TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
;  FILE REFERENCE: 1533.0790000
;  CURRENT APPLICATION NUMBER: US/09/677,575
;  CURRENT FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 09/220,081
;  PRIOR FILING DATE: 1998-12-23
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 1
;  LENGTH: 3621
;  TYPE: DNA
;  ORGANISM: Corynebacterium glutamicum
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (199)..(3621)
US-09-677-575-1

  Query Match       97.8%;   Score 3398.8;   DB 4;   Length 3621;
  Best Local Similarity  98.6%;   Pred. No. 0;
  Matches 3427;   Conservative 0;   Mismatches 47;   Indels 0;   Gaps 0;
```

```
Qy      1 GTGACTGCTATCACCCTGGCGGTCCTTGTTGTGAAAGGAATAATTACTCTAGTGTCGACT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 GTGACTGCTATCACCCTGGCGGTCCTTGTTGTGAAAGGAATAATTACTCTAGTGTCGACT  207

Qy     61 CACACATCTTCAACGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGGCGAAATC    120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 CACACATCTTCAACGCTTCAGCATTCAAAAGATCTTGGTAGCAACCGCGGCGAAATC    267

Qy    121 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCGAACCGGTAGCTATTTACCCC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCGAACCGGTAGCTATTTACCCC 327

Qy    181 CGTGAAGATCGGGGATCATTCCACCGCTCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    328 CGTGAAGATCGGGGATCATTCCACCGCTCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACC 387

Qy    241 GAAGGCTCACCAGTCACAAGGCGTACCTGGACATGAATGAAATTATCGGTGTCAGCTAAAAAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 GAAGGCTCACCAGTCACAAGGCGTACCTGGACATGAATGAAATTATCGGTGTCAGCTAAAAAA 447

Qy    301 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 GTTAAAGCAGATGCTATTTACCCGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCC   507

Qy    361 CGGAGTGCGCGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 CGGAGTGTGCGCGGAAAACGGCATTACTTTATTGGCCCAACCCCAGAGGTTCTTGATCTC  567

Qy    421 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAAGAAGGCGGTGGTCTGCCAGTTTTGGCG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568 ACCGGTGATAAGTCTCGTGCGGTAACCGCCGGAAGAAGGCGGTGGTCTGCCAGTTTTGGCG   627

Qy    481 GAATCCACCCGAGCAAAAACATGACAGATCGATGAGATCGTTAAAGGCGTAAAGGCCAGACTTAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    628 GAATCCACCCGAGCAAAAACATGACAGATCGATGAGATCGTTAAAGGCGTAAAGGCCAGACTAC  687

Qy    541 CCCATCTTTGTAAAGCGCCGCAAATTAGCAACAGAAGCATCTCGTGAAGCGTCAGATTCTTCA  600
            |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 CCCATCTTTGTGAAGCGCCGCAAATTAGCAACAGAAGCATCTCGTGAAGCGTCAGATTCTTCA   747

Qy    601 CCTGATGAGCTCCGCAAAATTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCGCCTTTCGGC 660
            |||||||||||||||  |||||||||||||||||||||||||||||||||||||||  |||||
Db    748 CCTGATGAGCTCCGCAAAATTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCGGCTTTCGGC  807

Qy    661 GACGGTTCGTATATGTCGAACGTGCTGTGATTAACCCCCAGCACATTGAAGTGCAGATC   720
            |||  || ||||||||||||||||||||||||||||| |||||  |||||||||||||||
Db    808 GATGGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATC    867

Qy    721 CTTGGCGATCGCACTGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT   780
            ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    868 CTTGGCGATCACACTGAGAAGTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGT   927
```

```
Qy   2521 CCAGGCCCAACCGGTCGCGTCTACCGCCGACAGAAATCCCAGGCGGACAGTTGTCCAACCTG 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2668 CCAGGCCCAACCGGTCGCGTCTACCGCCGACAGAAATCCCAGGCGGACAGTTGTCCAACCTG 2727

Qy   2581 CGTGCACAGGCACCGCACTGGCCTTGCTGTGATCGCTTCGAGCTCATCGATCGAAGACAACTAC 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2728 CGTGCACAGGCACCGCACTGGCCTTGCTGTGATCGCTTCGAGCTCATCGATCGAAGACAACTAC 2787

Qy   2641 GCAGCCGTTAATGAGATGTCTGGGACGCCCAACCAAGGTCACCCATCCTCCAAGGTTGTT   2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2788 GCAGCCGTTAATGAGATGTCTGGGACGCCCAACCAAGGTCACCCATCCTCCAAGGTTGTT   2847

Qy   2701 GGGGACCTGCACTCCACCTGGTTGCGGGTGTAGATCCAGCAGACTTTGCTGCGAGAC     2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2848 GGGGACCTGCACTCCACCTGGTTGCGGGTGTAGATCCAGCAGACTTTGCTGCGCCGAT    2907

Qy   2761 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGCTTCCTGCGCGGCGGGCGAGCTTGGTAAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2908 CCACAAAAGTACGACATCCCAGACTCTGTCATCGCGCTTCCTGCGCGGCGGGCGAGCTTGGTAAC 2967

Qy   2821 CCTCCAGGTGGCTGGCCAGAACCACTGCGCACCGCCGCACCCCGCACTGGAAGGCCCGCTCCGAAGGC 2880
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2968 CCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCGCCGCACCCCGCACTGGAAGGCCCGCTCCGAAGGC 3027

Qy   2881 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGGCCACCTCGACCGTGATGATTCC   2940
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2928 AAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGGCCACCTCGACCGTGATGATTCC   3087

Qy   2941 AAGGAACGTGCAACGCCTCAACCGCCTCTGCGTTCCCGAAGCCAACCAACGAGAGTTCCTC   3000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3088 AAGGAACGTCGCAATAGCCTCGAACCGCTGCTGTTCCCGAAGCCAACCAACGAGAGTTCCTC   3147

Qy   3001 GAGCACCGTCGCGCGCTTCGGCAACACCTCTGCGCTGACATGATCGGAATTCTTCACGGA    3060
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   3148 GAGCACCGTCGCGCGCTTCGGCAACACCTCTGCGCTGACATGTCGGAATTCTTCACGGC    3207

Qy   3061 CTGGTCGAGGGCCGAGACTTTGATCCGCTGCCAGATGTGCCGCACCCCACTGCTTGTT    3120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3208 CTGGTCGAGGGCCGAGACTTTGATCCGCTGCCAGATGTGCCGCACCCCACTGCTTGTT    3267

Qy   3121 CGCCTCGGATGCGATCTCTGAGCCAGACGATAAGGCGTATGGCTTCCTGTCCGCAACGTC   3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3268 CGCCTCGGATGCGATCTCTGAGCCAGACGATAAGGCGTATGGCTTGGTGCGGCCAACGTC   3327

Qy   3181 AACGGCCAGATCCGCCCGCCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3328 AACGGCCAGATCCGCCCGCCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCGCAACC 3387

Qy   3241 GCAGAAAAGGCAGATTCCTCCAACAAGGGCATGTTGCTGCGCACCATTCGCTGCGTTGTC   3300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3388 GCAGAAAAGGCAGATTCCTCCAACAAGGGCATGTTGCTGCGCACCATTCGCTGCGGTTGTC   3447

Qy   3301 ACTGTGAAGATGAAGCAACAACATACTGCTTCTCGTTGACGGCAAGATTGAACGCGTTGTG   3360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3448 ACCGTGAAGATGAAGCAACAACATACTGCTTCTCGTTGACGGCAAGATTGAACGCGTTGTG   3507

Qy   3361 GCTATGAAGATGGAAGCAACAACAATCACTGCTTCTCGTTGACGGACTTGATCGTCGTTCCTAA 3420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3508 GCTATGAAGATGGAAGCAACAACAATCACTGCTTCTCGTTGACGGACTTGATCGTCGTTCCTAA 3567

Qy   3421 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGGACTTGATCCGTCGTTCGTTCCTAA 3474
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3568 GTTCCTGCTGCAACGAAGGTGGAAGGTGGCGGACTTGATCCGTCGTTCGTTCCTAA 3621
```

```
RESULT 2
US-09-677-575-1
;  Sequence 1, Application US/09677575
;  Patent No. 6403351
;  GENERAL INFORMATION:
;  APPLICANT: Sinskey, Anthony J.
;  APPLICANT: Lessard, Philip A.
;  APPLICANT: Willis, Laura B.
```

```
Qy   781 CGTCACCAAAAGTGTCGAAATTGCGCCAGCAGCATTTGGATCGAGAACTGCTGAT   840
         |||||||||||||||||||||||||||||||||||||| |||||| ||||| ||||
Db   928 CGTCACCAAAAGTGTCGAAATTGCGCCAGCAGCACGAACTTTGGATCGCGTGAT   987

Qy   841 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 CGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACC  1047

Qy   901 GTGGAATTCTTGGTCGATGAAAAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATC   960
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||| ||||
Db  1048 GTGGAATTCTTGGTCGATGATGAAAAGGGCAACCACCAGTCTTCATCGAAATGAACCCACGTATC  1107

Qy   961 CAGGTTGAGCACACCGTGACTGACGAGGTCACCGAGGTGACCTGGTGAAGGCGCAGATG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108 CAGGTTGAGCACACCGTGACTGACGAGGTCACCGAGGTGACCTGGTGAAGGCGCAGATG  1167

Qy  1021 CGCTTGGCTCGTGTCGCAACTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1080
         ||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||
Db  1168 CGCTTGGCTCGTGTCGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACC  1227

Qy  1081 CACGGTGCAGCACTGCAGTCGCCGATCACCACCGGAAGATCAAACAACACGGCTTCCGCCCA  1140
         ||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||
Db  1228 CACGGTGCAGCACTGCAGTCGCCGATCACCACCGGAAGATCAAACAACAACGGCTTCCGCCCA  1287

Qy  1141 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT  1200
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  1288 GATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGT  1347

Qy  1201 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGTGAAATGACC  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 GCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGTGAAATGACC  1407

Qy  1261 TGCCGGTGGTTCCGACTTTGAAACTGCTGTGTTGCGTGTGCTGCAGCGCGCGTTGGCCGTGAGGAC  1320
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db  1408 TGCCGTGTCTGCGACTTTGAAACTGCTGTGTTGCGTGTGCTGCAGCGCGCGTTGGCCGTGAGGAC  1467

Qy  1321 ACCGTGTCTGGTGTTGCAACCAACATTGGTTCTTGCGTTCTGCGGTGCGTTGCGCGGGAAGAGGAC  1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468 ACCGTGTCTGGTGTTGCAACCAACATTGGTTCTTGCGTTCTGCGGTGCGTTGCGCGGGAAGAGGAC  1527

Qy  1381 TTCACTTCCAAGCAGATCTGCCATCGCCACCGGATTTATCGGCGATCCTGGCAGATTACTTGGCAGATGTCACCGTGAAC  1440
         |||||||||||||||| | ||||||||||||||||| |||||||||||||||||||||||||||||||
Db  1528 TTCACTTCCAAGCAGATCTGCCATCGCCACCGGATTCGGCGATCCTGGCAGATTACTTGGCAGATGTCACCGTGAAC  1587

Qy  1441 CCACCTGCCGGATGATGAGCAGAGGACGCCATCTGATTACTTGGCAGACCAACCTGGATTACCACACCTTCAGGCT  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1588 CCACCTGCCGGATGATGAGCAGAGGACGCCATCTGATTACTTGGCAGACCAACCTGGATTACCACACCTTCAGGCT  1647

Qy  1501 AAGCCTCATGGTGTGGCGTCCAAAGGATGTGCAGCACCAATCGATAAGCTGCCCAACATC  1560
         |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  1648 AAGCCTCATGGTGTGGCGTCCAAAGGATGTGCAGCACCAATCGATAAGCTGCCTAACATC  1707

Qy  1561 AAGGATCTGCCACTGCCACCGCGGTTCCACGGTTCCCGTGAGCAGCTTGGCCCAGCCGCG  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1708 AAGGATCTGCCACTGCCACCGCGGTTCCCGTGAGCAGCTTGGCCCAGCCGCG  1767

Qy  1621 TTTGCTCGTGATCTCCGTGAGCAGGCAGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT  1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1768 TTTGCTCGTGATCTCCGTGAGCAGGCAGACGCACTGGCAGTTACTGATACCACCTTCCGCGAT  1827

Qy  1681 GCACACCAGTCTGCTTGTGCCGACCCGAGTCGCTCGCACTGAAGCCTGCGGCGCAGAG  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1828 GCACACCAGTCTGCTTGTGCCGACCCGAGTCGCTCGCACTGAAGCCTGCGGCGCAGAG  1887

Qy  1741 GCCGTGCAAAGCTGACTCCTGAGCTTTGTCCGTGAGGCGGCGGCGGACCTAC  1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1888 GCCGTGCAAAGCTGACTCCTGAGCTTTGTCCGTGAGGCGGCGGCGGACCTAC  1947

Qy  1801 GATGTGGCGGATGCGTTTCCTCTTTGAGGAATCGTGGGACCAGGCTCGACGAGCTGCGCGAG  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1948 GATGTGGCGGATGCGTTTCCTCTTTGAGGATCGTGGGACCAGGCTCGACAGGCTCGACGAGCTGCGCGAG  2007
```

```
Qy  1861 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGATACACC  1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2008 GCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCCGCAACACCGTGGATACACC  2067

Qy  1921 CCGTACCCAGACTCCGTCCGTCCGCGCGGTTGTTAAGGAAGCTGCCAGCTGCTCCGGCGTGGAC  1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2068 CCGTACCCAGACTCCGTCCGTCCGCGCGGTTTGTTGTTAAGGAAGCTGCCAGCTGCTCCGGCGTGGAC  2127

Qy  1981 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA  2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2128 ATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCA  2187

Qy  2041 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT  2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2188 GTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTATGGCTTATTCTGGTGATCTCTCT  2247

Qy  2101 GATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGCGAGAGATCGTC  2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2248 GATCCAAATGAGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGCGAGAGATCGTC  2307

Qy  2161 AAGTCTGGCGTCACATTCTGGCCATTAAGGATATGGCTGGTCTCGTTGCCAGCTGCG  2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2308 AAGTCTGGCGTCACATTCTGGCCATTAAGGATATGGCTGGTCTCGTTGCCAGCTGCG  2367

Qy  2221 GTAACCAAGCTCGTCACCGCACTGCGCGTGGCGTGAATTCGATCTGCCAGTGCACGTGCACACC  2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2368 GTAACCAAGCTCGTCACCGCACTGCGCGTGGCGTGAATTCGATCTGCCAGTGCACGTGCACACC  2427

Qy  2281 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCGGTGGTGCAGAT  2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2428 CACGACACTGCGGGTGGCCAGTTGGCTACCTACTTTGCTGCAGCTCGGTGGTGCAGAT  2487

Qy  2341 GCTGTTGACGGTGCTTCCGGAGCTGGGAAGCTGTGCGCCGATACCGGTTTGACCTGCCAGCCATCCCAGCCTCCGTGTCTGCC  2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2488 GCTGTTGACGGTGCTTCCGGAGCTGGGAAGCTGTGCGCCGATACCGGTTTGAGCTTGAGCCATCCCAGCCTCCGTGTCTGCC  2547

Qy  2401 ATTGTTGCTGCAGCCGTACTGGGAAGCTGGGATCCAGGAAATCCCAGGCGGACAGTTGTCCAACCTG  2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2548 ATTGTTGCTGCAGCCGTACTGGGAAGCTGGGATCCAGGAAATCCCAGGCGGACAGTTGTCCAACCTG  2607

Qy  2461 GACCTCAGCCGTTACCGTCGGCGCAGCTGTGGCGCCCGCCCAGCGCGGACAGTCATCATCCGCTCATCGAGAACAACTAC  2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2608 GACCTCAGCCGTTACCGTCGGCGCAGCTGTGGCGCCCGCCCAGCGCGGACAGTCATCATCCGCTCATCGAGAACAACTAC  2667

Qy  2521 CCAGGCCCAACCCGTCCGTCTCAACCGCCTTGCTCGCGTCATCGCTTGCTGAGCTCATCCGAAGACAACCTG  2580
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2668 CCAGGCCCCAACCCGTCCGTCTCAACCGCCTTGCTCGCGTCATCGCTTGCTGAGCTCATCCGAAGACAACCTG  2727

Qy  2581 CGTGCACAGGCCACCACCGACACTGCGCTCGCTCATCGAGCTCATCGAGAACAACTAC  2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2728 CGTGCACAGGCCACCACCGACACTGGGCCACCGACTCGTTTCGAGTCATCATCGAGACTCATCGAGAACAACTAC  2787

Qy  2641 GCAGCCGTTAATGAGATGCTGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2788 GCAGCCGTTAATGAGATGCTGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTT  2847

Qy  2701 GGGGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCAGACTTTGCTGCGAGAC  2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2848 GGGGACCTCGCACTCCACCTGGTTGGTGCGGGTGTGGATCCAGACTTTGCTGCGCCGAT  2907

Qy  2761 CCACAAAGTACGACATCCCGAGACTCGTCCTGCGCGTTCCTGCGGCGAGCTTGGTAAC  2820
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2908 CCACAAAGTACGACATCCCAGACTCGTCCTGCGCGTTCCTGCGGCGAGCTTGGTAAC  2967

Qy  2821 CCTTCCAGGTGCTGGCCAGAACCACTGCGACCACTGCGGAAGGCCGCTCCGAAGGC  2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2968 CCTTCCAGGTGGCGGTGGCCAGAGAGCACTGCGACCACTGCGGAAGGCCGCTCCGAAGGC  3027

Qy  2881 AAGGAACCTGCGACGAAGAGTTCCTGACGCCACCTGACGATGATTCC  2940
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3028 AAGGAACCTGCGACGAAGAGTTCCTGACGCCACCTGACGATGATTCC  3087

Qy  2941 AAGGAACGTCGCAACAGCCTCAACCGCCTTGCTTCCCGGCAACCGCTCTGTTCCCGAAGAGTTCCTC  3000
```

```
Db   3088  AAGGAACGTCGCAATAGCCTCAACCCGTCTGCGCTGCTGTTCCCGAGCAGAGTTCCTC  3147
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Qy    264  CCTGGACATGCCAATATGCGGTCGCAGCTAAAAAGTTAAAGCAGATGCTATTTACCC   323

Db   3001  GAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATATCGTGAATTCTTCTACGGA  3060
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    204  TTTGAATATCGAACGTATCATCATCGAAACGCCGCGAGTAGTCTCTCGCGCAGGTCTC   263

Db   3148  GAGCACCGTCGCCGCTTCGGCAACACTCTCGCGTCGGATGATCGTGAATTCTTCTACGGC  3207
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    324  GGGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCGGAGTGCCGGAAAACGGCAT   383

Db   3061  CTGGTCGAGGGCGGAGACTTGATCGGTCCGAGATGTGGCGCACCCACTGCTTGTT  3120
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    264  TGGGTATGGTTTTTTAAGTGAAAAATGAACATTTGACGCGTGCTGCTGAGGAAGGCAT   323

Db   3208  CTGGTCGAAGGCCGGAGACTTTGATCCGGCTGCCACCCCACCGCTGCTTGTT  3267
           |||||||||||||||||||||||||||||||||||||||||||||||
Qy    384  TACTTTTATTGGCCAACCCAGAGGTTCTTGATCTGATCAGACATGTTTGGAGAATAGGGCTAG   443

Db   3121  CGCCTCGGATGCGGATCTCTGAGCCAGCAGATAAGGCTATGCCCAATCGGT  3180
           |||||||||||||||||||||||||||||||||||||||||||
Qy    324  TAAATTTATAGTCCGATCTTGAACATCTAGACATGTTTGGAGAATAGGGCTAG   383

Db   3268  CGCCTCGGATGCGATCTCTGAGCCAGCAGATAAGGCTATGCCCAATCGGT  3327
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    444  AACCGCCGGAAGAAGGCGCTGGTCTGCGACTTACCCCAGCAGTTTGGCGG---AATCCACCCCCAGCAAAAA   500

Db   3181  AACCGGCCAGATCCGCCCAATCGTGTCGTGTGCCGTGACCGGCTCCGTCGT  3240
           ||||||||||||||||||||||||||||||||||||||||||
Qy    384  AACAACTGCTATTAACGCTAACTTACCTGTAATCCCGGGTACAGATGTCCTATTGAAAG   443

Db   3328  AACGGCCAGATCCGCCCAATCGTGTCGTGTGCCGTGACCGGCTCCGTCGT  3387
           |||||||||||||||||||||||||||||||||||||||||||||
Qy    501  CATCGATGACATCGTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAAGGCAGT   560

Db   3241  GCAGAAAAGGCAGATTCCTCCAACAAGGGGCCATGTGCTGCTGCATCGCTGTGGTTGTGTC  3300
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    444  TTTTGAAGCTGCAGAACAGTTTGCTAATGAAGCAGTTACCCACTTAGGCCAC   503

Db   3388  GCAGAAAAGGCAGATTCCTCCAACAAGGGGCCATGTGCTGCCATCGCTGTGGTTGTGTC  3447
           |||||||||||||||||||||||||||||||||||||||||||||
Qy    561  TGCCGGTGGTGGCGGACGCGGTATGCGCTTGTTCTTCGTGATGAGCTCCGCAAATT   620

Db   3301  ACTGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTCAGTCGCAATCATCGAG  3360
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    504  AAGCGGTGGCGGTGGTAAAGTATGCGGAATCGTTCGTGAAGCTGAAGCGCATTCGGCGGCCATTAGAAGACGC   563

Db   3448  ACCGTGACTGTTGCTGAAGGTGATGAGGTCAAGGCTCAGTCGCAATCATCGAG  3507
           |||||||||||||||||||||||||||||||||||||||||||
Qy    621  GGCAACAGAGCATCTCGTGAAGCTGAAGCCGAATCGGCCACGCCGTTCGTATATGTCGA   680

Db   3361  GCTATGAAGATGGAAGCAACAATCACTCGTCTTCTGTTGACGGCAACAATCTGATCGCGGTTGTG  3420
           |||||||||||||||||||||||||||||||||||||||||||||
Qy    564  TTTCCATCGTGCGAAATCAGAAGCCGAAAAGTCATTTGGTAATAGCGAAGTTTATATCGA   623

Db   3508  GCTATGAAGATGGAAGCAACAACAATCACTCGTCTTCTGTTGACGGCAAAATCGATCGCGGTTGTG  3567
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    681  ACGTGCTGTGATTAACCCCCAGCATTGAGTGCAGCAGTCCTGGGATCGCGACTCGGAGA   740

Db   3421  GTTCCTGCTGCTGCAACGAAGGTGGAAGGTGGGCGACTTGATCGTCGTTCCTAA  3474
           |||||||||||||||||||||||||||||||||||||||||||||||
Qy    624  AAGATATATTGATAATTCCAAAGCATTATTGGTGATGAATTCGGGAA   683

Db   3568  GTTCCTGCTGCAACGAAGGTGGAAGGTGGGCGACTTGATCGTCGTTCCTAA  3621
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    741  AGTTGTACACCTTATGAACGTGACTGCTCACTGCAGCGGTCGTCACCAAAAGTTGTCGA   800

                                                                            Qy    684  TATCATTCATTTGTATGAAAGATTGCTCCGTACAACAGCAGTCATCAAAAGGTTGTTGA   743

RESULT 3                                                            Qy    801  AATTGCGCCAGCACAGCCATTTGGATCCAGATGCCTGATCGATCCAGAACTGCGTGTGCGGATGCAGT   860
US-09-134-001C-591                                                           ||||||||||||||||||||||||||||||||||||
; Sequence 591, Application US/09134001C                            Db    744  AGTTGCACCTTCAGTAGGTCTTTCAACAAATTAAGAGAGGGAATTTGTGATGCCGCAAT   803
; Patent No. 6380370
; GENERAL INFORMATION:                                             Qy    861  AAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGAATTCGTTGGAATTCTTGGTCGATGA   920
; APPLICANT: Lynn Doucette-Stamm et al                                      ||||||||||||||||||||||||
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS   Db    804  TCAACTGATGGAAAAATAAAATACGTCAACGCTGGAACAGTAGAATTTTTAGTTTCTG-   862
;       EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007                                          Qy    921  AAAGGGCAACCACGTTTCATCGGAAATGAACCCCACGTATCCAGGTTGAGCACACCGTGAC   980
; CURRENT APPLICATION NUMBER: US/09/134,001C                                |||||||||||||||||||||||||
; CURRENT FILING DATE: 1998-08-13                                  Db    863  --GGGATGAATTTTTCTTCATTGAGGTTAATCCACCGTGTTCAAGTTGAGCATACAATTAC   920
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08                                    Qy    981  TGAAGAAGTCACCGAGTGGACCTGGTGAAGGCGCAGATGCGGTTCGGCTGGTGGTCGCAAC  1040
; PRIOR APPLICATION NUMBER: US 60/055,779                                   |||||||||||||||||||||||||
; PRIOR FILING DATE: 1997-08-14                                    Db    981  TGAAATGATTACTGGTATAGACATTGTGAAAACGCAAATTTTTAGTTGCTAATGGAGAAATC   980
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 591                                                    Qy   1041  CTTGAAGGAATTGGGTCTGACC-------CAAGATAAGATCAACAGCACCACGGTGCAGC  1091
; LENGTH: 3465                                                              |||||||||||||||||||||||
; TYPE: DNA                                                        Db    981  GTTATTTGGAGATAAAATCTCTATGCCACAGATCCAGATTTATGATGATTTCTGACAGTATGC  1040
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-591                                                 Qy   1092  ACTGCAGTGCCGCATCCACCGGAGATCCCGAGATCAACAGCGCTTCCGCCCAGATACCCGAAC  1151

Query Match      13.7%;    Score 476.6; DB 4; Length 3465;         Db   1041  GATACAATGTCGTATATACACTGAAGATCCTACTAATGATTTTATGCCAGATTCTGCAC  1100
Best Local Similarity  48.2%;  Pred. No. 2e-122;
Matches 1655;  Conservative  0; Mismatches 1724;  Indels  54;  Gaps  9;    Qy   1152  TATCACCGGTACGCCTCACCAGGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCT  1211

                                                                   Db   1101  AATTATTGCATATGCTTGTTGCTAACCGTGGTGAAATCGCCATTAGGAGATTTTTAGAGCGGC  1160

Qy     84  ATTCAAAAAGATCTTGTAGCAACCGCGGCAAATCGCCGGTCCGCTTCCGTGTCAGCT   143   Qy   1212  C--GTGGCGAAATCACCGGCACTTGACTCGTGGTGAAAATGACCTGCCCTGG  1268
           ||||||||||||||||||||||||||||||||||||||
Db     84  AATAAAGAAAATTACTTGTTGCTAACCGTGGTGAAATCGCCATTAGGAGATTTTTAGAGCGGC   83   Db   1161  CCAAGGTGCAGAAATTTCACCTTACTCACGATTCACTCAGTTAGTTAAGCTTTCTACACATGC  1220

Qy    144  ACTCGAAACCGGTCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCA   203   Qy   1269  TTCCGACTTTGAAACTGCTTCTGGAAACTGCAGAGAAAATGGAACGTTCATTACGCGCAAATGCGAATTCG  1328
           ||||||||||||||||||||||
Db     24  AGCAGAATAAATATCAGTACAGTAGCAATTTATTCTAATGAAGATAAAAGTTCGTTACA   143   Db   1221  CGTTTCATTAAACAGCTGAAGAGAAATGAACGTCATTACGCGCAAATGCGAATTCG  1280

Qy    204  CCGCTCTTTGCTTCTGAAACTGTCCGCATTGGTACTGGACAGGGACCGTA   263

Db    144  TAGATATAAAGCAGATGAATCCTATCTCAGTTGGAAGTGATTTAGGACCTGCTGCTGAAAGTTA   203
```

us-09-974-973-1.rni

```
Db  2361  ATTATATTATGCACTAAATGGATTTCCACGTAATTAAGAACTGATATTGGGTTAGA  2420
                     ||||||| ||  ||  ||| ||||||||     ||||| ||||||   || ||
Qy  2421  TGACCTCGAGCCGTACTGGGAAGCTGTGCGCGGACTGTACCTGCCATTGAGTCTGGAAC  2519

Db  2421  AGAGTTGAGTCATTACTGGTCTGTAGTCAGACCTTACTATGCAGACTTTGAGAGTGATAT  2480
                     |||||| ||  |||||  ||| |||  |||||||||| ||||||||
Qy  2520  CCCAGGCCCAACGGTCGCGTCTACCGCCACGAAATCCCAGGCGACAGTTGTCCAACCT  2579

Db  2481  CAAATCACCAAATACAGAAATTTATCAACATGAAATGCCAGGTGGCCAAATATTCAAACTT  2540
                 |||| |||||||| ||||| |||||||||||
Qy  2580  GCGTGCACGGCCACCCGACTGGGCCTTGCTGCTGCTCATCGAAGACAACTA  2639

Db  2541  AAGTCAACAAGCTAAAAGTTTAGGATTGGGCGAACGTTTTGATGAAGTCAAAGAGATGTA  2600
                |||||||||| |||||||||||
Qy  2640  CGCAGCCGTTAATGAGATGCTGGGACCGCCAACCAAGGTCACCCCATCCTCCAAGGTTGT  2699

Db  2601  TCGTCGTGTCAACTTCCTGTTTGGAGATCTTGGTGCGGGTAGATCCAGCAGACTTTGCTGCAGA  2660
Qy  2700  TGGCGACCTCGCACTCCACCTGGTTGGTGCGGGTGTAGATCCGCGGCGGACGAGCTTGGTAA  2759

Db  2661  TGGAGATATGGCACTATATATGGTGCAAAATGATCTTGATGAGAATACGGTCATCAATGA  2720
Qy  2760  CCCACAAAAGTACGACATCCCAGACTCGTCATCGCGGTCCTCGCGCGGCGAGCTTGGTAA  2819

Db  2721  TGGTTATAAATTAGAGTTTCCCAGAAATCTTGTTGTCATTCTTTAAAGGTGACATTGGACA  2780
Qy  2820  CCCTCCAGGTGGCTGGCCAGAACCACTGCCGCACCGCACTGGAAGGCCGCTCCGAAGG  2879

Db  2781  ACCTGTCAACGGATTCAACAAGAAATTGCAAGATGTTATTTATTTAAAGGACAGCAACCAAT  2840
Qy  2880  CAAGCGACCTCTGACGGAAGTTCCTGAGAAGAGCAGGCGCACCTGAGCGCTGATGATTC  2939

Db  2841  TACTGAAAGACCAGGTGAATACTTGGAGCCGGTCGAATTTTGAAGCAATCCGTCAAGAATT  2900
Qy  2940  CAAGA--------------ACGTCGCAACAAGCCTCAACCGCCTGCTTGTCTCC  2978

Db  2901  AAGCGGACATACACAGAGGTAACAGAACAAGATAATAATTAGTATGTATGTACTTTATCC  2960
Qy  2979  GAAGCCAACCGAGAGTTCCTCGAGCAGCTGCCAGCTCGGCGCGCTCTGGA  3038

Db  2961  GAAGGTATATAAACATAATATATTCAAACGAAAGACAATTGGTAATGTATCTTTACTGGA  3020
Qy  3039  TGATCGTGAATTCTTCTACGGACTGGTCGTGAGGGCCGCCGAGGACTTTGATCCGCCTGCCAGA  3098

Db  3021  TACACCCGACATTCTTATTTGGCATGCGTAATG---GTGAAACAGTTGAAATTGAAATTGA  3077
Qy  3099  TGTGCGCACCCCACTGCTTGTTCCGCCTGATGCCGATCTCTGAGCGCAGACGATAAGGGTAT  3158

Db  3078  TACTGTAAACGTCTAATTATTAAATTATAAGGCGATAAATCAAATCAAATCATATTGG  3137
Qy  3159  GCGCAATGTTGTGGCCAACGTCACGCCCAATGGTCAACGGCCAGATCTGTGCGTGACCGCTC  3218

Db  3138  GCGTACAATTTATTACGCTATGAATGGTCAAGCAAGACGTATCTATATTCAAGATGAAAA  3197
Qy  3219  CGTTGAGTCTGTCACCGCACCCGATTCCTCCAACAAGGGCCATGTTGC  3278

Db  3198  TGTTAAAACGAATGCTAAAGTTAAACCTAAGGCCGGATAAATCAAATCAAATCATATTGG  3257
Qy  3279  TGCACCCATTCGCTGGTGTTGTCACT---GTGACTGTTGCTGAAGGTGATGAGGTCAAGGC  3335

Db  3258  TGCTCAAATGCCTGGTTCTGTAACTGAAGTCAAAGTGTCTGTAGGCCGGCGATGGAAGTTCAAGC  3317
Qy  3336  TGGAGATGCCAGTCGCCAATCATCGAGCTATGAAGATGCAACAATCACTGCTTCTGT  3395

Db  3318  TAATCAGCCATTATTAATCACTGAAGCAATGACGATTCAGGCGACCATT  3377
Qy  3396  TGACGGCAAGATTGAACGCGTTGTGTTCCTGCAACGAAGGTGGCGACTT  3455

Db  3378  TGATGGAATTATTAAAACAAATCAATGTTGCTAATGGAGATGCCATTGCCACAGGAGATTT  3437
Qy  3456  GATCGTCGTCGTT  3468
```

```
Qy  1329  TGGTGTTGCAACAACATTGGTTCTTGCGTGCGTTGCTGCGGGAAGAGGACTTCACTTC  1388
                |||||  ||||||||||||||| ||||| ||||||||| |||||||||||
Db  1281  TGGCGTAAAGACGAATATTCCATTTCTCATCAATGTTATGCGTAATGATAAAATTTAGAAG  1340

Qy  1389  CAAGCGCATCGCCGGATTTATCGGCGATCACCCACCTCCTTCAGGCTCCACCTGC  1448
                |||||||||||| ||||||| ||||| |||||| |||||  ||| ||| ||||
Db  1341  TGGTGATTATACTACTAAATTTATTGAAGAAACACCTGAACTTTTCGATAATTGCACCGAC  1400

Qy  1449  GGATGATGAGGAGGAGCGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCA  1508
                ||| ||||||||| || || ||||| ||||| |||||| || |||| |||||
Db  1401  ATTGGACACGAGTACCAAGACTTTAGAGTATTGTAATGCACGAATCGATAAAACGGATTCC  1460

Qy  1509  TGGTGT------GGGTCCAAAGGATGTTGCAGCACCAATCGATAAGCTGCCCAACATC--  1560
                 ||| ||
Db  1461  TAATGTAGAAAAGCGTCCACTGCCACCGCCGTTCCCGTGACCGCCTTGGCCCAGCCGC  1520

Qy  1561  -AAGGATCTGCCACTGCCACGCGGTTCCCGTGAGCAGCTTGGCTACTGATACCACCTTCCGGCA  1619
Db  1521  AAAGAAAATCAATCAGTTATTTGGAACAAAACAAATTCTTGAGCAACATGGACCAACAGG  1580

Qy  1620  GTTTGCTCGTGATCTCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGGCGA  1679
Db  1581  TGTTACAAAATTGGGTTAGAGAACAAGAGATGTTTTAATTACCGATTACTACATTTAGAGA  1640

Qy  1680  TGCACACCAGTCTTTGCTGTTGCGACCCGAGTCCGTCATTCGGCTCTGAGAGCCTGCGGCAGA  1739
Db  1641  TGCACACAATCTTTACTTGCAACACGTGTAAGAACAAAAGATATGATGAACATTGCATC  1700

Qy  1740  GGCCGTCGCAAAGCTGACTCCTGAGCTTTGTCCGTGGAGGCCTGGGCGCGGCGACCTA  1799
Db  1701  TAAAACTGCTGAAGTTTTTAAAGATAGTTTTCATTGAGAAATGTGGGGTGTGCAACATT  1760

Qy  1800  CGATGTGGCGATGCGTTCCTCTTTGAGATCCGTGGACAGGCTCGACCAGCTGCGCGA  1859
Db  1761  TGATGTGCCCTATAATTTCTTGAAAGAGAATTCCATGGAACGTTTAGAAAGATTGCGCAA  1820

Qy  1860  GGCGATGCGAATGTAAACATTCAGATGTCTGCTTCGCGGCCGCAACACCCTGGGATACAC  1919
Db  1821  AGCCATTCGATGTGTTATTCCAAATGTTATTACCGAGCTTCAGCGCAGTTCCGGCGAGTTTATATA  1880

Qy  1920  CCCGTACCCAGACTCCGTCTGCCGCGCGGTTTGTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGA  1979
Db  1881  AAACTATCCTGATAATGTAATTAAGAAATTCGTTCATGGAAAGTGCAAAAGCTCGGTGTAGA  1940

Qy  1980  CATCTTCCGCATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGC  2039
Db  1941  TGTTTTCCGTATATTCGACTCATTGAACTGGGTTGATCAAATGAAAGTAGCGAATGAAGC  2000

Qy  2040  AGTCCTGGAGACCAACACCGCCGGTAGCCGAGGTGGCTATGCGTTATTCTGGTGATCTCTC  2099
Db  2001  TGTTCAAGAAGCTGGAATGGTATCTGAGGGTACAAATTGCTATACAGGTGATATTTTAA  2060

Qy  2100  TGATCCAAATGAAAAGCTCTACACCCTGGATTTACTACCTAAAGATGGCAGAGGAGATCGT  2159
Db  2061  TGCTCAAAACGTTCCAATATTTATACTTTAGATTACGTTAAATGCTAAAGAACTGGA  2120

Qy  2160  CAAGTCTCTGGAGACCAACACCGCCGGTAGCCATTCTCTGCCATTAAGGATATGGCTCGTC  2219
Db  2121  AAGAGAAGATTCCATTATTAGCAATTAAAGATATGGCTGCTTGGTTTATTGAAACCGAAAGC  2180

Qy  2220  GGTAACCAAGCTGCTTCCCGCACTCGCTGTCTGCGACCACCTCCCAGCCATCCTGCACAC  2279
Db  2181  AGCTTACGGAATTAAATTGGTGAATTACGTGAGGCCAACACATCTTCCAATTCATTTACATAC  2240

Qy  2280  CCACGACACTGCCGGGTGGCCAGTTGGCTACCTACTTGCTGCGAGCTCAAGCTGGTGCAGA  2339
Db  2241  ACATGATACTAGTGGAAATGGATTGTTGGACATATAAACAAGCAAATTGATGCTGGGTAGA  2300

Qy  2340  TGTCTGTTGACGGTGCTTCCCGGCACCACTGCTCTGCGACCACCTCCCAGCCATCCCTGTCTGC  2399
Db  2301  TATTATAGATACTGCTGTTGCATCTATGAGTGGTTAACGAGTCAACAAGTGCAAATTC  2360

Qy  2400  CATTGTTGCTGCATTCGCGATACCCGTCGCGATACCCCGTTGAGCCTCGAGGCTGTTTC  2459
```

Db   3438 ATTAGTGGAAATT 3450

RESULT 4
US-07-956-700B-5
; Sequence 5, Application US/07956700B
; Patent No. 5539092
; GENERAL INFORMATION:
;   APPLICANT: Robert Haselkorn and Piotr Gornicki
;   TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
;   TITLE OF INVENTION: Carboxylase
;   NUMBER OF SEQUENCES: 116
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: 321 No. 5539092th Clark Street
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60610
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy Disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII-DOS
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/956,700B
;   FILING DATE: 19921002
;   CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Thomas E. No. 5539092thrup
;   REGISTRATION NUMBER: 33,268
;   REFERENCE/DOCKET NUMBER: ARCD:058
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 1-312-744-0090
;   TELEFAX: 1-312-755-4489
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1362 base pairs
;   TYPE: Nucleic acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Linear
;   MOLECULE TYPE: Oligonucleotide
US-07-956-700B-5

Query Match   8.8%;   Score 306.6;   DB 1;   Length 1362;
Best Local Similarity   53.6%;   Pred. No. 2.3e-75;
Matches   714;   Conservative   0;   Mismatches   604;   Indels   15;   Gaps   3;

Qy   85   TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAAATGCGCGGTCCGGTCTTTCCGTGCAGCA 144
        ||||| ||||||||||| ||||| ||||| ||| ||| ||| |||
Db   7   TTCAACAAGATCCTGATCGCGGCCAATCGCGGCGAAATGCCCTGCCATTCTCCGCACTTGT 66

Qy   145   CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCAC 204
        |||| || ||| ||||| |||| |||| || ||||||| ||| |||
Db   67   CAAGAACTCGGGATCGGCAGCAGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy   205   CGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAAGGCTCACCAGTCAAGGCGTAC 264
        ||| |||| ||| |||||| ||||| ||||
Db   127   GTGCAGTTAGCGGACGAAGCGGCGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy   265   CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAAGTTAAAGCAGATGCTATTTACCCG 324
        ||| || ||||| ||| |||||| ||| |||||| ||| |||
Db   184   CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCTAATGCCAGCGCCATTCACCCC 243

Qy   325   GGATATGGCTTCCTGCTTCTGAAAATGCCGGAGTGCGCGGAAAACGGCATT 384
        || ||||| || |||||| |||||| |||| |||||
Db   244   GGCTATGGCTTCTTGGCGGAGAATGCCCGCCTTTGCAGAAATCTCGGCCGATCACCCATCTC 303

Qy   385   ACTTTTATTGGCCAACCCCAGAGGTTCTTGATCTCACCGTGATAAGTCTCGTGC---G 441
        || ||||||||| |||| |||||| || ||| ||||||| |||
Db   304   ACCTTTATTGGCCCCAGATTCGATTCGAGCCATGGGCGATAAATCCACCGGTAAG 363

Qy   442   GTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTTGCGGAATCCACCCCCGAGCAAAAAC 501

Db   364   GAAACAATGCAGCGGTTCCGGACGATTCCGGCGCAGTAGCGCTCGCTGACGGAT 423
        |||| ||| || ||| ||| ||| || ||
Qy   502   ATCGATGACATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAGGCAGTT 561

Db   424   GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTCATCCCGTCATGATCAAAGCGACG 483
        || ||| ||| ||| ||| ||| || ||||| || ||
Qy   562   GCCGGTTGGTGGCGGACGGGTATGGCGTTGGTTCTTCTTCACCTGATGAGCTCCGCAAATTG 621

Db   484   GCGGGGGCGGTGGTCGCGGATGCCGGTCTGCAGATCTGAAAAAACTG 543
        ||| ||| || ||| || || |||
Qy   622   GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCGAA 681

Db   544   TTCCTTGCTGCCAAGGAGGAGCCGAGGCAGCTTTGGAATCCAGGACTGTATCTCGAA 603
        |||| || || ||| |||
Qy   682   CGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCGACTGAGAA 741

Db   604   AAATTTATGATCGCCCACGCCACGTTGAATTTCAGATTCTTGGCCTACGGCAAT 663
        |||| || |||| ||| ||| ||||
Qy   742   GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTGCTGTCAGCAAAAGTGTCGAA 801

Db   664   GTAGTGCATCTAGGCGAGCGGCGATTCGTCCATTCAACGTCGTCACCAAAAGCTGCTCTGAA 723
        |||||| ||| ||| ||| || |||
Qy   802   ATTGCGCAGCACAGCATTTGGATCCAGAACTGCGTGATCGCATTGTGCGGATGCAGTA 861

Db   724   GAAGCCCCAGTCCGGCATCGGCAGACCTGCGGCCAGAAAAATGGGCGATGCCGCCCTC 783
        ||| |||
Qy   862   AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGAACCGTGGAATTCTTGGTCGATGAA 921

Db   784   AAAGTGCTCAAGCGATCGGCTACATCGGTGCCGGCCACATCGGTACTGGGTTTCTGGTCGATGCG 843
        ||| ||| || || ||| |||
Qy   922   AAGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCAGTTGAGCACACCGTGACT 981

Db   844   ACCGGCAACTTCTACTTCATGGAGATGAATACCCGATCCAAGTCAAGTCGAGCATCCAGTCACA 903
        |||
Qy   982   GAAGAAGTCACCGAGGTGGACCTGGTGAAGGGCAGATGCGCTTGGCTGCTGTGTCAACC 1041

Db   904   GAAATGATTACCGGACTGGACTTGATTGCGGACAGATTCGGATTGCCCAAGGCG----- 958
        |||
Qy   1042   TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACTCAACTGCGCGGCGATCGAATGC 1101

Db   959   ----AAGCGCTGCGCTTCCGGCAAGCGCGATATTCAACTGCGGGGCCCAGATACCACCGCG 1014
        |||
Qy   1102   CGCATCACCACCACGGAGATCCAAACAACACGGCTTCCGCCCAGATACCACCGCCG 1161

Db   1015   CGTATCAATGCGGAGAGTCCGGAATACAATTTCCGGCCGAATCCTGGCCGCCATTACAGGC 1074
        ||| ||| ||| |||
Qy   1162   TACCGCTCACCAGGCCGAGCTGGCGTTCGTCTTGACGGTCGAGCTCAGCTCCGGTGGGAA 1221

Db   1075   TATTTACCGGCCCCGGCCGCCCCGGCGAAATCGCGTTCGTGTCGATTCCCATGTTTATACCGACTACGAA 1134
        |||
Qy   1222   ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTTTGAA 1281

Db   1135   ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTGTCGGGGTGCAACACGGGAA 1194
        |||
Qy   1282   ACTGCTGTTGCTGCGGGCGCGGTGGTGCGTCGTGTTGCAACC 1341

Db   1195   GAGGCGGATCGCGCGGATGCAGTGCGTGCTCTGCGCGGGAAGGCGATGGCCATCACCGGCATCGCC 1254
        |||
Qy   1342   AACATTGGTTTCTTGCTGCGGAGTTCATCAGCTAGCTAGTTGCAGATGCCTCAAGCGCATCGCC 1401

Db   1255   ACCCTTAGTTCATCACGCCCTAATGCCACGCGCCATTCACCCC 1314
        |||
Qy   1402   ACCGGATTTATCG 1414

Db   1315   ACCAACTTTGTTG 1327

RESULT 5
US-08-476-537-5
; Sequence 5, Application US/08476537
; Patent No. 5756290
; GENERAL INFORMATION:

```
APPLICANT:  Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION:  Carboxylase
NUMBER OF SEQUENCES:  116
CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Arnold, White & Durkee
    STREET:  321 No. 5756290th Clark Street
    CITY:  Chicago
    STATE:  Illinois
    COUNTRY:  USA
    ZIP:  60610
COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy Disk
    COMPUTER:  IBM PC Compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  ASCII-DOS
CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/476,537
    FILING DATE:
CLASSIFICATION:  435
PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  07/956,700
    FILING DATE:  10/21/92
ATTORNEY/AGENT INFORMATION:
    NAME:  Thomas E. No. 5756290thrup
    REGISTRATION NUMBER:  33,268
    REFERENCE/DOCKET NUMBER:  ARCD:058
TELECOMMUNICATION INFORMATION:
    TELEPHONE:  1-312-744-0090
    TELEFAX:  1-312-755-4489
INFORMATION FOR SEQ ID NO:  5:
SEQUENCE CHARACTERISTICS:
    LENGTH:  1362 base pairs
    TYPE:  Nucleic acid
    STRANDEDNESS:  Single
    TOPOLOGY:  Linear
    MOLECULE TYPE:  Oligonucleotide
US-08-476-537-5
```

```
Query Match     8.8%;     Score 306.6;  DB 1;   Length 1362;
Best Local Similarity  53.6%;  Pred. No. 2.3e-75;
Matches  714;  Conservative  0;  Mismatches  604;  Indels  15;  Gaps  3;

Qy    85  TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGGTCCTTCCGTGCAGCA  144
              ||||||| ||| |||  |  |||  || ||| || ||||| |||| |||| |||
Db     7  TTCAACAAGATCCTGATCGCCAATCGCGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT   66

Qy   145  CTCGAAACCGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCAATTCCAC  204
              ||| ||  | ||| ||| ||  |||| || ||||| ||||| || ||| ||||
Db    67  CAAGAACTCGGATGCCGACGATCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT   126

Qy   205  CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGCTGTCACCAGTCAAGGCGTAC  264
              | |||||||| |  |||||| | |||| ||||||| |||| |
Db   127  GTGCAGTTAGCGGACGAAGCGGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT  183

Qy   265  CTGGACATCGATGAAATTATCGGTGCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCCG  324
              | ||  || |||| |||  |||||| || || |||  || |||||| ||||||||| |||
Db   184  CTCAATATCCCCAACATCATTGCCGGCGCCCTGACCCCTAATGCCAGCGCCATTCACCCCC  243

Qy   325  GGATATGGCTTCCGTCTGAAAAATGCCCAGCTTGCCCCGCGAGTGCGCGGAAAACGGCATT  384
              |||| ||||||| ||  || |||  | ||||| |||  | |||  ||||||  |
Db   244  GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTGCGGAGAATCGCGCCGATCCACATCTC  303

Qy   385  ACTTTTATTGGCCAACCCAGAGTTCTGATCTCACCGGTGATAAGTCTCGTGC---G  441
              ||||| || ||  |||||| |||| |  | |||||||||  || ||  |||
Db   304  ACCTTTATTGGCCCCAGCCCGATTCGATTCGAGCCATGGGCGATAAATCCACCGCTAAG  363

Qy   442  GTAACCGCCGCGAAGAAGCGGGTCGGCGTTTGGCGGAATCCACCCCGAGCAAAAAC  501
              | ||  | ||  ||  || || || ||  ||| ||   |||||| ||| ||
Db   364  GAAACAATGCAGCGGTCGGCGTTCCGACGATTCCGGGCCAGTGACGGTCTGCTGACCGAT  423

Qy   502  ATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAAGGCAGTT  561
              |  |||  || || |  | |||  ||||  | |||||| || ||
```

```
Db   424  GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCATCAAAGGCGACG   483
              |  |||| || ||||||| ||||||||||||||||||||||||| | |  |  ||||
Qy   562  GCCGGTGGTGCGGCAGCGGGTATGCCGCTTTGTTTTCTTCCACTGATGAGCTCCGCAAATTG   621
              |  ||  || |||  ||  ||||| ||| ||| | ||||| ||  |  ||||||
Db   484  GCGGGGGGCGGGTGGTGTGGTGTATGCGGTGTGCGGTCCGGTGACCCTGCAGATCTGGAAAAACTG   543
              |  | ||| ||  ||| | ||  |  |||| ||| ||  ||||||||| || ||
Qy   622  GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGAGCCGAGCTTTTGGGAATCCGGAA   681
              |   ||| || || ||||||  |  ||||| || |||||  || ||
Db   544  TTCCTTGCTGCCCAAGGAGAAGCCGAGCCAGCTTTTGGGAATCCAAGACTGTATCTCGAA   603
              ||  ||||  ||||||  | | ||||| ||||  ||||
Qy   682  CGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCACTGGAGAA   741
              || |||| |  ||| | ||||  |  |||  ||  |||| ||||  ||||| |
Db   604  AAATTTATCGATCGGCCCCACGCGTTGAATTTCAGATCTTGGCCGATGCCTACGGCCAAT   663
              |||  ||  || | ||  |||  ||| |  ||| ||||  ||||| || |||
Qy   742  GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTGTCACAAAAAGTTGTCGAA   801
              || ||  || || || || ||  | ||| ||  ||||  || ||||||| || ||
Db   664  GTAGTGCATCTAGGCGAGCGGGATTGCTCCATTCAACGTCGTCACCAAAAGCTGCTGAA   723
              || || || | || ||  ||| |||  || || ||||| || || ||  || || |||
Qy   802  ATTGGCCAGCACAGCATTTGGATCCAGAACTGCGTCATGCAGTTGTGCGGATGCAGTA   861
              |||| |||||| |  ||| |||| |||  ||||| |  |  ||  || || ||  |
Db   724  GAAGCCCCAGTCCGGCGCTATCGGCAGAACCTGCGGCGAGAAAATGGGCGCGATGCCGCGCGTC   783
              | ||| || ||||  |  || ||  | ||| |||  |  |||  | ||| || |||  ||
Qy   862  AAGTTCTGCGCCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA   921
              |||||||||  || |||||  ||||  |||||| ||||  || || ||||||  |||||
Db   784  AAAGTCGCTCAAGCGATCGGCTACATCGGTGCACATCCGCACCGTGGAGTTTCGTGGTGGATGCG   843
              |||||||| |||| ||  ||| |  |||| ||| | ||||| |  || ||  || |||||
Qy   922  AAGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT   981
              ||||  || || ||| || |||  |||||| || ||  ||||||| ||  || ||| ||
Db   844  ACCGCAACTTCTACTTCATGAGGATGAATAACCCGCATCCCAAGTCGAGCATCCAGTCAGTCACA   903
              || || || |  || || |  |||| |||| ||||||  |||  |||||  || || ||
Qy   982  GAAGAAGTCACCGAGGTGACCTGGTGAAGGCGCAGATGCGCTTGGCTGCTGGTGCAACC  1041
              ||||||||  ||||||||  ||||  ||||||||||||||  ||||| |  | ||||
Db   904  GAAATGATTACGGACTGGACTTGATTGCGGAGCAGATTCGAGCAGATTGCCCAAGGCG-----   958
              |||  ||  |||||| ||  || || ||||| |||| | |  ||||  ||| |||
Qy  1042  TTGAAGGAATTGGGTCTGACCAAGATAAGATCAAGACCCACGTGCAGCACTGCAGTGC  1101
              ||| ||||  ||||  || || ||||| ||  |||  || |  |||||| || || ||
Db   959  ----AAGCGCTGCGCCTTCCGACCAGCCGATATTCAACTGCGCGGCGATCGAATGC  1014
              ||| || || || ||  ||||  | | ||  |  |||||| |||| ||  |
Qy  1102  CGCATCACCACGGAAGATCGAAGATCCAAACAACGCTCCGCCCAGATACCGGAACTATCACCGCG  1161
              || |||  ||||| ||||  | || |||  ||  || ||||||||  | || ||||||||||
Db  1015  CGTATCAATGCGGAAGATCCGGAATACAAATTCCGGGCCGAATCAATTTCTGACGGTGCCATTACAGGC  1074
              || |||  ||| ||||||||  |  |||| |  |  ||| ||| || ||  || || || ||||
Qy  1162  TACCGGCTCACCAGGCGGAGCTGGCGTTCGTTCGTGTCGATTCCATGTTTATACCGACTACGAA  1221
              || ||||| || || || ||| || || |||  ||||  ||| |||| || ||  || ||||
Db  1075  TATTTACCGCCCGGCGCCCGGCGAAATCGCGGTCGTGTGTGAAAATGACCTGCCCGGTTCCGACTTTGAA  1134
              |||| || || ||  ||  |  |||||  || || || || ||| || ||  || ||||||
Qy  1222  ATCACCGCACACTTTGACTCCATGCTGCTGAAGATCGGGGATCATTCCAC  1281
              || ||||||  | ||  |||| || |  |  |||  |||| || |||  ||
Db  1135  ATTCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA  1194
              || |||||  || || ||| ||| ||  ||||  || || || ||||  ||  ||||
Qy  1282  ACTGCTGTGTGCTTCGTGCACAGCGGCGGTTGGCTGAGTTCACCGTGTCGTGTTGTCAACC  1341
              || || || |  || |  ||  |  || |||  || |  ||| || |  || |  ||||
Db  1195  GAGGGCGATCGCGCGATGGAAGCGGTGCTCTGCGGAATGCGCCATCACCGGCTTGCCCGACG  1254
              || || |||  || || |  || |  || || ||  ||| || ||  |  |||  ||| ||
Qy  1342  AACATTGGTTTCTTGGTGCGGTTGCAGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGGAACTCTAT  1401
              |||||||  |  ||  || || |  || || ||  |  ||| |||||  |  || || ||  ||
Db  1255  ACCCTTAGTTTCCATCAGCTGATGTTGCAGATCGAGTTCGTGCGCGGCGGGAACTCTAT  1314
              ||||  ||||  |||||| | ||| || |||| | |||| || || || ||| ||||||
Qy  1402  ACCGGATTTATCG  1414
              ||||  ||| |||
Db  1315  ACCAACTTGTTG  1327
```

```
RESULT 6
US-08-485-607-5
; Sequence 5, Application US/0848S607
; Patent No. 5792627
; GENERAL INFORMATION:
;     APPLICANT:  Robert Haselkorn and Piotr Gornicki
;     TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;     TITLE OF INVENTION:  Carboxylase
;     NUMBER OF SEQUENCES:  116
;     CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: Arnold, White & Durkee
; STREET: 321 No. 5792627th Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,607
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/956,700
; FILING DATE: 10/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas E. No. 5792627thrup
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: ARCD:058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-312-744-0090
; TELEFAX: 1-312-755-4489
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1362 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Oligonucleotide
US-08-485-607-5

Query Match   8.8%;   Score 306.6; DB 1; Length 1362;
Best Local Similarity 53.6%; Pred. No. 2.3e-75;
Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

Qy   85 TTCAAAAAGATCTTGGTAGCAAACCGGCGAAATCGCGGCTCCGTGCTTTCGGTGCAGCA 144
        |||| || |||| | | |||||||||||||||||||| |||||| |||| | ||
Db   7 TTCAACAAGATCCTGATCGGCAATCGCGGCGAAATCGCCCTGCCGCATTCTCCGCACTTGT 66

Qy 145 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGAATCGGGATCATTCCAC 204
        |||||| || ||||| ||| || ||| |||||||||| ||||| |||||||||||||
Db   67 CAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACCTGGGATCGCAACGCGCTCCAT 126

Qy 205 CGCTCCTTTGCTTCTGAAGCTGTCCCGCATTGGTACTGTGAAGGCTCAAGGCGTAC 264
        |||| ||| |||||||||||| | ||| ||| || || |||||||| |||
Db 127 GTTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy 265 CTGGACATGATGAAATTATCGGTGCGAGCTAAAAAGTTAAAGCAGATGCTATTTACCCG 324
        ||| || | | || | ||| ||| ||||||||||||| | | |||| ||||||
Db 184 CTCAATATCCCCAACATCATTGCGGCCCTGACCCCTAATGCCACGGCCATTCACCCC 243

Qy 325 GGATATGGCTTCCTGTCTGAAAAATGCCCAGAGGTTCTTGATCTTCACCGTTGATAAGTCTCGTGC---G 384
        |||||||||||| | | |||||||| | | || | || | ||||| | ||| | ||||| |
Db 244 GGCTATGGCTTCTTCTGCGGAGAATGCCCGCGTTTGCGGAATCCACCCCGAGCCAAAAC 303

Qy 385 ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 441
        || ||||||||||| | ||| | || | || | |||| | ||||| | ||| |
Db 304 ACCTTTATTGGCCCCCAGCCCGATTCGAGCCCATGGGGCGATAAATCCACCCCGCTAAG 363

Qy 442 GTAACCGCCGCAGAAGAGGCTCGGCGTTGGTCTGCCCAGTTTGGCGGAAT 501
        |||| | || |||||||| | || ||| | | ||||| |||||||
Db 364 GAAACAATGCAACATCGTTAAAAGCGGCTGAAGGCCCAGACTTACCCCATCTTTGTAAAGGCAGTT 423

Qy 502 ATCGATGACATCGTGCCCAAGTTGCTGCCGAGATCGGCTATCCCGTCGTCTTTGTAAAGGCAGTT 561
        ||||||||||||||| ||||||||||||||||||||||||||||||||| | | | ||| |||
Db 424 GTTGATTCGGCTGCCAAGTGTGCTGCCGAGATCGGCTATCCCGTCGTCATGATCAAAGCGACG 483

Qy 562 GCCGGTGGTGCCGGACGGCGTGGTCGCGGTATGCGGTATGCGGT 621
        || |||||| || || | || |||| ||||||| || ||||||
Db 484 GCGGGGGGCGGCCGGTGTGTCGCGGTATGCGGGCCTGCGGTCGGTGACCCTGCCGAGATCTGGAAAAAACTG 543

Qy 622 GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCCGAA 681
        |||| || || |||| ||||||||||||| || ||||| | | || || |||||||||
Db 544 TTCCTTGCTGCCCAAGGAAGAGGCCGAGGCAGCTCTTTGGGAATCCAGGGACTGTATCTGAA 603

Qy 682 CGTGCTGATGAATAACCCCCAGCACAATTGAAGTGCAGAATCCTTGGCCGATCGCACTGGAGAA 741
        |||| || || |||| ||| | | || |||| | | ||| | | ||| ||||||||
Db 604 AAATTATCGATCGCCCACGTTGAGATTTCAGATTTCGGTCACCGATGACCTCGTACGGCCAAT 663

Qy 742 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCACAAAAAGTTGTCGAA 801
        | |||||| || | ||| | || |||| ||| ||| |||| ||| ||| | |||||
Db 664 GTAGTGCATCTAGGCGGCGGCGATTGCTCCATTCCATTCAACGTCGCCCAAAAGCTGCTGAA 723

Qy 802 ATTGCGCCAGCACAGCATTTGGATCCAGAACTTGCGTCGATCGGCTGCATTTGTGCGGATGCAGTA 861
        || | || || | | ||| | | ||| | | | |||| ||||| | |||| ||| |
Db 724 GAAGCCCCAGTCCGGCGATCTGCGGGCAGACCTTGCCGGCAGAAAATGGGGCGATGGCCGGCCGTC 783

Qy 862 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGGCGGAACCGTGAATTCTTGGTCGATGAA 921
        ||||| || | || || | | | | | |||| | | | || || | | ||
Db 784 AAAGTGCCTCAAGCGATCGGCTACATCGGTCGCGGCTACATCCGGTTCGTGGTCGATGCG 843

Qy 922 AAGGGCAACCACGTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 981
        || | |||| || | ||| | | | | | | || | |||| | || | |
Db 844 ACCGGCCAACTTCTACTTCATGAGATGAATAACCCGCATCCAAGTCAAGTACCAGTCACA 903

Qy 982 GAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCCAGATGCCTGGCTGCTGTCGGTGCAACC 1041
        |||||| | | | | ||||| || | |||| | || || || | ||| | || |
Db 904 GAAATGATTACGGGACTTGAGTTGATTGCGGAGCAGATTCAACTCGATCGGATCGAATGC------ 958

Qy 1042 TTGAAGGAATTGGGTCGCGCTTCCGGCAAGCCGATATTCAACTGCCCGGCGCACTGCAGTGC 1101
        | || || || | ||| | || ||||| | || | | | |||||||| |||||
Db 959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATCAACTGCCCGGCGCACTGCAGTGC 1014

Qy 1102 CGCATCACCACGGAGAATCAAAACACAGCGCTTCCGCCCAGATACCGGAACTATCACCGCG 1161
        || ||| | | | | | ||| | || || | | || ||| | || || ||||
Db 1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGCCGCCATTACAGGGC 1074

Qy 1162 TACCGCTCACCGGCGAGCTGGCGTTCGTCGTTCCGTGGGCCGAA 1221
        ||| | || ||| || | ||| | ||||||| | || |||
Db 1075 TATTTACCGCCCGGCGGCGGGCCGGTTCCCATGTTTATACCGACTACGAA 1134

Qy 1222 ATCACCGCACACTTGACTCCATGCTGGTGAAAATGACCTCCGACTTTGAA 1281
        ||| ||| | | || | | || || || | ||| ||| ||||||
Db 1135 ATTCCCGCCTATTACGGATTCGCTGATTGGTGCAAATTGATTGTGTCTGGGGTGCAACACGGGAA 1194

Qy 1282 ACTGCTGTGTCCGTGCACAGCGCCGCGTTGGCTGAGTTCACCGGTCGTGTTGCAACC 1341
        || ||||| | ||| |||| | | | || | || | | | | ||||| |||||
Db 1195 GAGGGCGATCGCGGCGGATGCGCAGCGGTGCTCTGCGGGGAATGCGGCCATCACCGCCGTTGCCCACG 1254

Qy 1342 AACATTGGTTTCTTGCGTGCCGTTGCTGCGGGGAAGAGGACTTCACTTCACGCATCGCC 1401
        |||| |||| | || ||| | | || | || |||||| | | |||| ||| |||
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGGTTCAGATGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy 1402 ACCGGATTTATCG 1414
        |||| ||| |||
Db 1315 ACCAACTTTGTTG 1327

RESULT 7
US-08-475-879-5
; Sequence 5, Application US/08475879
; Patent No. 5972644
; Patent No. 5972644 5786170
GENERAL INFORMATION:
;   APPLICANT: Robert Haselkorn and Piotr Gornicki
;   TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
;   TITLE OF INVENTION: Carboxylase
;   NUMBER OF SEQUENCES: 116
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: 321 No. 5972644 5786170th Clark Street
;   CITY: Chicago
;   STATE: Illinois

; COUNTRY: USA
; ZIP: 60610
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,879
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/956,700
; FILING DATE: 10/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas E. No. 5972644 5786170thrup
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: ARCD:058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 1-312-744-0090
; TELEFAX: 1-312-755-4489
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1362 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Oligonucleotide
US-08-475-879-5

Query Match   8.8%;   Score 306.6;   DB 2;   Length 1362;
Best Local Similarity 53.6%;   Pred. No. 2.3e-75;
Matches 714; Conservative 0; Mismatches 604; Indels 15; Gaps 3;

Qy   85 TTCAAAAAGATCTTGGTGCTAGCAAACCGCGGCGCAAAATCGCGGTCCGTGCTTTCCGTGCAGCA 144
             ||||||| |||||||||||| ||||||||||||||||
Db   7 TTCAACAAGATCCGATCGCCAATCGCGGCAAATCGCGGCGAAATCCCGCGCCCACTTGT   66

Qy 145 CTCGAAACCGGTGCAGCCACGGTAGCTATTACCCCCGTGAAGATCGGGATCATTCCAC 204
             || |||||||||||||||||| |||||||||
Db   67 CAAGAACTCGGGATCGGCACGAGTCGCCGTTCACTCCACTGTGGATCGCAACGCGCTCCAT 126

Qy 205 CGCTCTTTGCTTCTGAAGCTGTCCGCATTGTACTGTGAAGGCTCAAGGCGTAC 264
             |||||| |||| |||
Db 127 GTGCAGTTAGCGGACGAACGGCGTCTGTATTGG---CGAAGCGGCCAGCAGCAAAAGCTAT 183

Qy 265 CTGGACATGATGAAATTATCGGTGCGAGCTAAAAAAGTTAAAGACAGATGATGTATTACCCG 324
             || |||| |||||||||
Db 184 CTCAATATCCCCAACATCATTGCGGCGGCCCTGACCCCTAATGCCAGCGCCATTCACCCC 243

Qy 325 GGATAATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGGGAAAATCGGCATT 384
             ||||||||||||||||||
Db 244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy 385 ACTTTTATTGGCCAACCCAGAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 441
             ||||||||||||
Db 304 ACTTTTATTGGCCCCAGCCCCGATTCGATTCGAGCCATGGCGAATAAAATCCACCGCTAAG 363

Qy 442 GTAACCGCCGCGAAGAGAGGCTCGGTCTGCCAGTTTGGCGGAATCCACCCCGAGCAAAAAC 501
             ||||||||||||||
Db 364 GAAACAATGCAGCGGGTCGGCGGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy 502 ATCGATGACATCGTTAAAAGCGCTGAAGGCCAGAGACTTACCCCATCTTTGTAAAGGCCAGTT 561
             ||||||| || ||
Db 424 GTTGATTCGGCTGCCAAAGTTGCTGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy 562 GCCGGTGGTGGGGACGCGGTATGCGTGCGGGTATGCCGGTGGTCGCGATTCGGCGACAGGT 621
             |||||
Db 484 GCGGGGGGCGGTGGTCGCGGTATGCCGGTATGCCAGTCTGCAGATCTGGAAAAAACTG 543

Qy 622 GCAACAGAAGCATCTCGTGAAGCTGGGAAGCGGCATTCGGCGACGCGGTTCGGTATATGTCGAA 681
             ||||||||||
Db 544 TTCCTTGCTGCCCCAAGGAGAGAAGCCGGAGCCAGCGTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy 682 CGTGCTCTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGCACTGCTGAGGAA 741
             ||||||||||||||
Db 604 AAATTTATCGATCGCCCACGCCACGTTGAATTTCAGATTCTTGGCCGATGCCTACGGCAAT 663

Qy 742 GTTGTACACCTTTATGAACGTGACTCGCTCACTGCACGCGTCGTCACCAAAAAGTTGTCGAA 801
             |||||||||||||||
Db 664 GTAGTGCATCTAGGCGAGGCGATTGCTCATTCAACGTCGTCACCAAAAGCTGCTCGAA 723

Qy 802 ATTGCGCGACCACAGCATTTGGATCCAGAACTGCGTGATCGCATTGTGCGGATGCGAGTA 861
             ||||||||||||
Db 724 GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAGAAAATGGGCGGATGCCGCCGTC 783

Qy 862 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 921
             || |||||||||||||
Db 784 AAAGTCGCTCAAGCGATCGGCTACATCGGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy 922 AAGGGCAACCACGTTTCATCGAAAATGAACCCACGTATCCAGGTTGAGCACCACCGTGACT 981
             | ||||
Db 844 ACCGGCACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGAGCATCCAGTCACA 903

Qy 982 GAAGAAGTCACCGAGGTGGACCTGGTGAAGGGCGCAGATGCGCGTGCTGGCTGCGTCGTGCAACC 1041
             || |||||||||
Db 904 GAAATGATTTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTCGGCCCAAGGCG------ 958

Qy 1042 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1101
             |||||||||||
Db 959 ---AAGCGCTGCGCTTCCGCAAGCCGGATATTCAACTGCCAAGGCTTCACTCGCGATCGGATGC 1014

Qy 1102 CGCATCACCACGGAAGATCCAAACAACAACGGCTTCGCCCCAGATACCGGAACTATCACCCG 1161
             |||||||||||||
Db 1015 CGTATCAATGCGGAAGATCGGGAATACAATTTCCGGCCGAATCCTGCCGAATCCGCATTACAGGC 1074

Qy 1162 TACCGCTCACCAGGCGAGCTGGCGTTCGTCTTGACGTGCAGCTCAGTCGGTGGCGAA 1221
             ||||||||||||
Db 1075 TATTTACCGCCGCCCGGCGGCCCCGGCGTTCCGTGTCGTGAAAATGACCTGCGCGTGGTTCCGACTTTGAA 1134

Qy 1222 ATCACCGCACACTTGACTCCATGCGTCGTGGATCGGGATCATTCCAC 1281
             |||||||||
Db 1135 ATTCCGCCCTATTTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGGAA 1194

Qy 1282 ACTGCTGTTGTCGTCGTGCACAGCCGCGGCGGCCGTTGGCTGAGTTCACCGTGTCGTTGTGCAACC 1341
             ||||||||||||
Db 1195 GAGGCGATCGCCGCGGATGCGGGATGCGGCTGCTCGCGGGAAGTGGGCAATGCGCCATCACCGGACG 1254

Qy 1342 AACATTGGTTCTTGCGTGCGTTGCTGCGTGCGGGAGGACTTCACTTCCAAGCGCCATCGCC 1401
             ||||||||||
Db 1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCCTGCGCGGGGGAACTCTAT 1314

Qy 1402 ACCGGATTTATCG 1414
             ||| ||||||||
Db 1315 ACCAACTTTGTTG 1327

RESULT 8
US-09-433-043B-5
; Sequence 5, Application US/09433043B
; Patent No. 6399342
; GENERAL INFORMATION:
; APPLICANT: HASELKORN, ROBERT
; APPLICANT: GORNICKI, PIOTR
; TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-CoA CARBOXYLASE
; FILE REFERENCE: ARCD:338US
; CURRENT APPLICATION NUMBER: US/09/433,043B
; CURRENT FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 08/475,879
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/956,700
; PRIOR FILING DATE: 1992-10-02
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1362

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-433-043B-5

Best Local Similarity   53.6%;   Pred. No. 2.3e-75;
Query Match       8.8%;   Score 306.6;   DB 4;   Length 1362;
Matches   714;   Conservative   0;   Mismatches 604;   Indels   15;   Gaps   3;

Qy   85  TTCAAAAGATCTTGGTAGCAAACGCGGTCCGTTCCGTTCGGTGCAGCA 144
             ||||| |||||||||||| ||||||||||||||| |||||||||||||
Db    7  TTCAACAAGATCCTGATCGGCCAATCGGCGGCAAATCGCCCTGCCAATTCCCGCACTTGT 66

Qy  145  CTCGAAACCGGTGCAGCCACGCGGTAGCTATTACCCCGTGAAGATCGGGATCATTCCAC 204
             ||| |||||||||||||||||||||||||| ||||||||  |||||| |||  ||||
Db   67  CAAGAACTCGGATCGGCACGATCGCCGTTCACTCACGTGGATCGCAACGCGTCCAT 126

Qy  205  CGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTGGATGCTCACCAGTCAAGGCGTAC 264
            |  |||||||||| |||| ||||||||| |||||||||| |||||||||||||||||
Db  127  GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAAGGCGGCCAGCAGCAAAGCTAT 183

Qy  265  CTGGACAATGCATGATGAAAATTATCGGTGCAGCTAAAAAGTTAAAGCACAGATGCTATTTACCCG 324
             ||||||||||| |||||| ||||| ||||||||||||||||||| |||||||||||||
Db  184  CTCAATATCCCCAACATCATTGCGCGAGCCCGACCCCTAATGCCGAGTCGAAATCTGCCGGATCACCATCTC 243

Qy  325  GGATATGGCTTCCTGTCTGAAAAATGCCCAGCTTGCCCGCGAGTGCGCGGAAAACGGCATT 384
             ||||||||||||||||| |||||||||||||||||||| |||||||| ||||||| ||
Db  244  GGCTATGGCTTCTTGGCGGAGAAATGCCCGGCTTTGCGGCGGATAGTCTCGTGC---G 303

Qy  385  ACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 441
             |||||||||||||||| |||||||||||||||||||||||||||||||||
Db  304  ACCTTTATTGGCCCCGATTCGATTCGAGCCCGATAAATCCACCGGTAAG 363

Qy  442  GTAACCGGCGGGAAGAAGGCTGGTCTGCCAGTTTGGCGGAATCACCCCCGAGCAAAAAC 501
            ||  |||||||||||||||| |||||||||| |||||||||||||||||||
Db  364  GAAACAATGCAGCGGGTCGGCGCGTTCCGACGATTCCCGGGCAGTGACGGCTGCTGACGGAT 423

Qy  502  ATCGATGACATCGTTAAAAGCGTGAAGGCCGAGACTTACCCCATCTTTGTAAAGGCAGTT 561
             ||||||  |||||||||||||||||||||||||||||||||||||
Db  424  GTTGATTCGGCTGCCAAAGTTGCTGCGCCGAGATCGGCTATCCCGTCATGATCAAAGCGACG 483

Qy  562  GCCGGTGGTGCGCGACGCGGTATGCGCGTTGTTTCTTCACCTGATGAGCTCCGCAAATTG 621
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  484  GCGGGGGGCCGGTGGTCCGGTATGCGGCTGCGTCCGAGATCTTGGAAAAACTG 543

Qy  622  GCAACAGAAGCATCTCGTGAAGCTGAAGCGGAGGCCATTCGGCGACGGTTCGGTATATGTCGAA 681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  544  TTCCTTGCTGCCCAAGGAGAAGGCGAGGCCAGCTTTTGGGAATCCAGGACTGTATCCTCCGAA 603

Qy  682  CGTGCTGTGATTAACCCCCAGAGTGCGCACATTGAAGTGCGACATCCTTGGCCGATCGCACTGAGAA 741
            ||||||||||| ||||||||||||||||||||||||||||||||
Db  604  AAATTTATCGATCGCCCACCACCGCCACGTTGAATTCAGATCTTGGCCCATGCCTACGGCAAT 663

Qy  742  GTTGTACACCTTTATGAACGTGACTGCTCACTCGCAGCGTCGTCACCAAAAAGTTGTCGAA 801
             |||||||||||||||||||||||||||||||||||||||||||||||
Db  664  GTAGTGCATCTAGGCGAGCCGCGATTGCTCAACGTCGTCACCAACGTGCTCGAA 723

Qy  802  ATTGCCGCAGCACAGCAGCATTGGATCCAGAACTGCGTGATCGCATTGTGCCGGATCGAGTA 861
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  724  GAAGCCCCCAGTCGCGCGTATCGGCGCAGACCTGCGCGCAGAAAATGGGGCGATGCCGCCGTC 783

Qy  862  AAGTTCTGCCGCTCCATTGGTTACCAGGGGCGGGAACCGTGGAATTCTTGGTCGATGAA 921
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  784  AAAGTCGCTCAAGCAGTCCTCAAGCCATCGGTGTGCCGGCACCGTGGAGTTTCTGGTCGATGCG 843

Qy  922  AAGGGCAACCACCACGTTTCATCGAAATGAACCCAGTAGCCTTGAGCACACCGTGACT 981
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844  ACCGGCAACTTCTCTACTTCATGGAGATGAAATACCCCATCCAAGTCGAGTCCAGTCACA 903

Qy  982  GAAGAAGTCACCGAGTGGACCTGGTGAAGCCGCAGATGCGCAGTTGGGTGCTGGTGCAAACC 1041
             ||||| ||||||||| |||||||  ||| ||||||||||||||||||||
Db  904  GAAATGATTACGGGGACTGGACTTGATTGCGGAGCAGCAGATTCCGATTCGCCTGCTTGCCCAAGGGCG----- 958
```

```
Qy 1042  TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCCACGGTGCAGCACTGCAGTGC 1101
                      |||| ||||||||||||||  ||||||||||| ||||||||||||
Db  959  ---AAGCGCTGCGCGTTCCGGCGAAGCCGATATTCAACTGCGGCGGCCATCGGCGGATATCACCGCG 1014

Qy 1102  CGCATCACCACGGAGATCCAAAACACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1161
             |||||||||| ||||||||||||||||  |||||||||||| |||||||||||||
Db 1015  CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCGGAATCCTGGCCGCCATTACAGGC 1074

Qy 1162  TACCGGCTCACCAGCCGGAGCTGGCGTTCGTCTTGACGGTGCGAGCTCAGCGTGGCGGAA 1221
             |||||||| |||||||||||||||||||||||||||||||| |||||||||||||||
Db 1075  TATTTACCGCGCCCGGCGCCCCCGGCGTTCGTGTGTCGATTCCCATGTTTATACCGGACTAGAA 1134

Qy 1222  ATCACCCGCACACTTTGACTCCATGCTGGTCGGTGAAAATGACCTGCCGGTGGTTCCGACTTTGAA 1281
             |||||||||||||||| ||||| |||||||||||||||||||||||| ||||||||
Db 1135  ATTCCGCCCATTACGATTCGCTGATTGGCAAATTGATTGTGTCGGGGTGCAACACGGGAA 1194

Qy 1282  ACTGCTGTTGCTGTGCACAGCGCGCGTTGGCGGTGAGTTGACCGTGTCGGTGTTGCAACC 1341
             |||||| |||||||||||  ||||||||||||||||||||||||||||||||||
Db 1195  GAGGCGATCGGCCGGATGCGAGCGTGCTTCTGCGGGGGAATGCGCCATCACCGGCTTGCCGACG 1254

Qy 1342  AACATTGGTTTCTTGCGTGCCTGCGTCGTGATGTTGCAGATGCGCCTGAGTTCCTGCGGGAACTCTAT 1401
             ||||||| ||||||| |||||||| ||||||||||||||||||||||||||||
Db 1255  ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCGCCTGAGTTCCTGCGCGGGGAACTCTAT 1314

Qy 1402  ACCGGATTATCG 1414
             |||||||||||
Db 1315  ACCAACTTTGTTG 1327
```

```
RESULT 9
US-08-611-107-7
; Sequence 7, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
;   APPLICANT: Haselkorn, Robert
;   APPLICANT: Gornicki, Piotr
;   TITLE OF INVENTION:   NUCLEIC ACID COMPOSITIONS ENCODING
;   TITLE OF INVENTION:   ACETYL-CoA CARBOXYLASE AND USES
;   TITLE OF INVENTION:   THEREFOR
;   NUMBER OF SEQUENCES: 40
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Arnold, White & Durkee
;       STREET: P.O. Box 4433
;       CITY: Houston
;       STATE:   Texas
;       COUNTRY: United States of America
;       ZIP: 77210
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/611,107
;       FILING DATE: Concurrently Herewith
;       CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US SN 07/956,700
;       FILING DATE: 02-OCT-1992
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US PCT/US93/09340
;       FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US SN 08/422,560
;       FILING DATE: 14-APR-1995
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Kitchell, Barbara S.
;       REGISTRATION NUMBER: 33,928
;       REFERENCE/DOCKET NUMBER: ARCD:221
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (512) 418-3000
;       TELEFAX: (512) 474-7577
```

```
;        TOPOLOGY:  linear
US-08-422-560A-7

Query Match    8.8%;    Score 305;  DB 2;  Length 1362;
Best Local Similarity   53.5%;  Pred. No. 6.4e-75;
Matches   713;  Conservative   0;  Mismatches  605;  Indels   15;  Gaps   3;
```

```
Qy       85  TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGGTCCTTCCGTCGCAGCA 144
             ||||| |||| | |||||||| |  ||||| |||||||| ||||  ||||||||||||||||
Db        7  TTCAACAAGATCCTGATCGCCGCACGGCGAAATCGCCCTGCGCATTCTCCGCACTTGT 66

Qy      145  CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCAC 204
             ||  |||||||| ||||| ||||||| ||||||||||||||||| ||||||||||  |||
Db       67  GAAGAACTCGGATCGGACGAAGCCGGTTCACTCCACTGTGGATCGCAACGCGGCTCCAT 126

Qy      205  CGCTCTTTGCTTCTGAAGCTGTCCGCATTGGTACTGGAAGGCTCACCAGTCAAGGCGTAC 264
             ||||| ||||  ||||||||||||||||||| || |||||| ||||| ||||||||||
Db      127  GTGCAGTTAGCGGACGAAGCCGGTCTGTATTGG---CGAAGCGGGCAGCAGCAAAAGCTAT 183

Qy      265  CTGGACAATGATGAAATTATCGGTGCAGCTAAAAAAGTAAAGCAGAATGCTATTTACCCG 324
             ||||| |||||||||| || ||||||||||| ||||||||||| |||| ||||||||||
Db      184  CTCAATATCCCAACATCATGCGACGGCGTCTGTATTGCCAGCGCCATTCACCCCC 243

Qy      325  GGATATGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGCGCGGAAAACGGCATT 384
             ||| ||||||||||  ||||||||||||| |||||||||||||| ||||||||||||||
Db      244  GGCTATGGCTTCTTGGCGGAGAATGCCCGCCTTGCGAGAAATCTGGGCCGCGAAATCTC 303

Qy      385  ACTTTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAGTCTCGTGC---G 441
             ||| ||| ||||| |||||||||| ||| |||||||||||||| || ||||  ||
Db      304  ACCTTTATTGCCCCCAGCCCGATTCGAGCCATGGCCATGGCGATAAATCCACCCGCTAAG 363

Qy      442  GTAACCGGCCCGGAAGAAGGCTGGTCTGCCAGTTTGGCGAATCCACCCCGAGCAAAAAC 501
             |   ||||||||| |||||| |
Db      364  GAAACAATGCAGCGGGTCGGCGTTCCGACGATTCCGGGCAGTGACGGTCTGCTGACGGAT 423

Qy      502  ATCGATGACATCGTTAAAGGCTGAAGGCCAGACTTACCCCCATCTTTGTAAAGGCAGTT 561
             ||||||||  | |||
Db      424  GTTGATTCGGTCGCCAAAGTTGCTGCGCCGGAGATCGGCTATCCCGTCATGATGAAGCGACG 483

Qy      562  GCCGGTGGTGGCGGACGCGGGTATGCGGGCTTTGTTCTTCACCTGATGAGCTCCGCAAATTG 621
             || ||| ||||| |||| |||
Db      484  GCGGGGGGCCGGTGGTGCGGGTATGCGGCTGTGCGCTGTCCGTGGAAAAAACTG 543

Qy      622  GCAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCGACGGTTCGGTATATGTCGAA 681
             ||||||| || ||  | ||||||
Db      544  TTCCTTGCTGCCCAAGGAGGAAGCCGAGGCAGCTTTTGGGAATCCACGGACTGTATCCGAA 603

Qy      682  CGTGCTGTGATTAACCCCAGCACATTGAAGTGCACATGCGGCGATCGCACTGACGAGAA 741
             ||||
Db      604  AAATTTATCGATCGCCACCCGCCACGTTGAATTCAGATTCTTGGCCGATGCCTACGGCAAT 663

Qy      742  GTTGTACACCTTTATGAACGTGACTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 801
             || |||
Db      664  GTAGTTGCATCTAGGCGAGCGCGATTGCTCCATTCAACGTGTCGTCACCAAAAAGCTGCTCGAA 723

Qy      802  ATTGCGCCAGCACAGCATTGGATCCAGAACTGCGTGATCCGATTGTGCCGGATGCAGTA 861
             ||||||
Db      724  GAAGCCCCCAGTCCGGCGCTATCGGCAGACCTGCGGCAGAAAATGGGCGATGCCCGGTC 783

Qy      862  AAGTTCGCGCCTCCATTGGTTACCAGGGGCGGGAACCGTGGAATTCTTGGTCACCGGTGACT 921
             ||||
Db      784  AAAGTTCGCTCAAGCGATCGGCTACATCGGTGCACGGCGCGCTGGAGTTTCTGGTCGATGCG 843

Qy      922  AAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGGTGACT 981
             || |
Db      844  ACCGGCAACTTCTACTTCATGGAGATGAATACCCGCATCCAAGTCGCATCCAGTCACA 903

Qy      982  GAAGAAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTCGTGCTGGTGCAACC 1041
             ||||
Db      904  GAAATGATTACGGACGCTGGACTTGATTGCCGGAGCAGATTCGGATTGCCCAGGCG----- 958

Qy     1042  TTGAGGGAATTGGGTCTGACCCAAGATAAGACAGAAGATCAAGACCCACCAGTGCCAGTAGTGC 1101
             ||||
```

```
Db      959  ----AAGGCTGCGCTTCCGGCAAGCCAAGCCGATATTCAACTGCCGCCAGATACCGGAACTATCACCGCG 1014

Qy     1102  CGGATCACCACGGAAGATCCAAACAACGGCTTCGCGCCAGATACCGGCCATTACCGG 1161
             ||||| |   ||||||||
Db     1015  CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCGAATCCATTACAGGC 1074

Qy     1162  TACCGCTCACCAGGCCGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAA 1221
             ||  ||||||
Db     1075  TATTTACCGCCCGGCCGCGCCCCCGGTTCGTGTCGTGTCGTGAAAAATGACCTCGCCGTGGTTCCGACTACGAA 1134

Qy     1222  ATCACCGCACACTTGACTCCATGCTGGTCGAAAAATGACCTGCCGGTGTTCCGACTTTGAA 1281
             ||||
Db     1135  ATTCCCGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTGGGGTGCAACACGGAA 1194

Qy     1282  ACTGCTCGTTGCTGCGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCGGTGTTGCAACC 1341
             ||  ||||
Db     1195  GAGGCGGATCGGCGGATGCAGCGTGCTCTGCGCGGGATGCGGGGGGAAAGAGGGACTTCCAAGCGGCATCGCC 1254

Qy     1342  AACATTGGTTTCTTGCGTGCGTTGCTGCGCGGAAGAGGACTTCACTTCCAAGCGCATCGCC 1401
             || |
Db     1255  ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCCTGAGTTCTGCGCGGGGGGAACTCTAT 1314

Qy     1402  ACCGGATTATCG 1414
             ||||
Db     1315  ACCAACTTTGTTG 1327
```

```
RESULT 11
US-08-468-793-7
; Sequence 7, Application US/08468793
; Patent No. 6177267
; GENERAL INFORMATION:
    APPLICANT: Haselkorn, Robert
    APPLICANT: Gornicki, Piotr
    TITLE OF INVENTION:  ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
    TITLE OF INVENTION:  METHODS OF USE
    NUMBER OF SEQUENCES: 29
    CORRESPONDENCE ADDRESS:
       ADDRESSEE: Arnold, White & Durkee
       STREET:  P.O. Box 4433
       CITY:  Houston
       STATE:  Texas
       COUNTRY:  United States of America
       ZIP:  77210
    COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS/ASCII
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/468,793
       FILING DATE:  06-JUN-1995
       CLASSIFICATION:  800
    PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US 08/422,560
       FILING DATE:  14-APR-1995
       APPLICATION NUMBER:  US SN 07/956,700
       FILING DATE:  02-OCT-1992
       CLASSIFICATION:  800
       APPLICATION NUMBER:  PCT/US93/09340
       FILING DATE:  30-SEP-1993
       CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
       NAME:  Kitchell, Barbara S.
       REGISTRATION NUMBER:  33,928
       REFERENCE/DOCKET NUMBER:  ARCD:152/KIT
    TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (512) 418-3000
       TELEFAX:  (713) 789-2679
       TELEX:  79-0924
; INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
       LENGTH:  1362 base pairs
```

```
;    TYPE: nucleic acid
;    STRANDEDNESS: single
;    TOPOLOGY: linear
;    MOLECULE TYPE: other nucleic acid
US-08-468-793-7

Query Match    8.8%;    Score 305; DB 4; Length 1362;
Best Local Similarity   53.5%;    Pred. No. 6.4e-75;
Matches  713; Conservative  0; Mismatches  605; Indels  15; Gaps  3;

Qy     85 TTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGGCGTCCGTGCTTTCCGTGCGAGCA 144
          |||||| ||||| || ||||||||||||||||||||||||||||||||| ||||||||||
Db      7 TTCAACAAGATCCTGATCGCAATGCGGCGGAAATCGCCCTGCGCCATTCTCCGCACTTGT 66

Qy    145 CTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGATCGGGGATCATTCCAC 204
          ||||| ||||||||||||||||||||||| |||||||||||||||||||||
Db     67 GAAGAACTCGGGATCGGCACGATCGCCGTTCACTCCACTGTGATCGCAACGCGCTCCAT 126

Qy    205 CGCTCTTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGGCTCACCAGTCAAGGCGTAC 264
          |||| | ||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    127 GTGCAGTTAGCGGACGAAGCGGTCTGTATTGG---CGAGCGCGCCAGCAGCAAAAGCTAT 183

Qy    265 CTGGACATCGATGAAATTATCGGTCAGCTAAAAAAGTTAAAGCAGATGCTATTTACCCG 324
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 CTCAATATCCCCAACATCATTGCGGCGAATGCCCGCGTAATGCCAGCGCCATTCACCCC 243

Qy    325 GGATATGATGGCTTCCTGTCTGAAAAATGCCCAGCTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 384
          |||| ||| |||||||||||||||||||||||||| ||||||||||
Db    244 GGCTATGGCTTCTTGGCGGAGAATGCCCGCTTTGCAGAAATCTGCGCCGATCACCATCTC 303

Qy    385 ACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGC---G 441
          |||| |||||||||||||||||| ||||||||||||||||||||||||||||||
Db    304 ACCTTTATTGCGCCCCGATTCGATTCGAGCCATGGGCGGATAAATCCACCGCTAAG 363

Qy    442 GTAACCGCCCGGAAGAAGGCTGGTCTGCGCCGTTCCGACGTTCGGCGGTGGCGGTCAGTT 501
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GAAACAATGCACGCGGGTCGGCGTTCCGACGATTCCGGCAGTGACGGTCTGCTGACGGAT 423

Qy    502 ATCGATGACATCGTTAAAAGCGCTGAAGGCCAGACTTACCCATCTTTGTAAAGGCAGTT 561
          |||| ||||||| || ||||||||||||||||||||||||||||| || ||||||||
Db    424 GTTGATTCGGCTGCCAAAGTTGCTGCGCGAGATCGCCGATATCCCGTCATCAAAGCGACG 483

Qy    562 GCCGGTGGTGGCGGACGCGGTATGCGCGCGTTTGTTCTTCACCTGATGAGCTCCGCAAATTG 621
          ||||||||| ||||||||||||||||||||||| || ||||||||||||||
Db    484 GCGGGGGGCGGTGTCGGGTATGCCGGTGGTCGGGGTCGGTGAGCGGCGTGCTGCGAAAAACTG 543

Qy    622 GCAACAGAAGCAATCTCTGAAGCGTGAAGCGGCATTCGGGCGACGTTCGGTATATGTCGAA 681
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    544 TTCCTTGCTGCCGCCAAGGAGGAGCCGAGCCAGCGCTTTTGGGAATCCAGGACTGTATCTCGAA 603

Qy    682 CGTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATGCCACTGGAGAA 741
          ||| |||||||||||| |||||||| ||||||||||||||||||||||||||
Db    604 AAATTTATCGATCGCCCAGCCACGTTGAATTCAGATCTTGGCCGATGCCTACGGCAAT 663

Qy    742 GTTGTACACCTTTATGAAGCGGACTGTCACTCGCAGCGTCGTCACCAAAAAGTTGTCGAA 801
          |||||||||| ||| |||||||||||||||||||||| || ||||||||||
Db    664 GTAGTGCCATCTAGGCGAGCGGCCGATTGCTTCCATTCAACGTGCTCACCACCAAAGCTGTCTTCGAA 723

Qy    802 ATTGCCGCCAGCACAGCAGCATTTGGATCCAGAACTGCCTGATCGCATTTGTGCCGGATGCAGTA 861
          ||| |||||||||||||||||||||||| ||||||||||||||||
Db    724 GAAGCCCCCAGTCCGGCGCTATCGCGGCGCGAACCTGCGGCGCAGAAAATGGGCGATGCCGCCGTC 783

Qy    862 AAGTTCTGCCGCCTCCATTGGTTACCAGGGCGCGGAATTCTTGGTCGATGAA 921
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    784 AAAGTCGCTCAAGCGATCGGCTACATCGGTCCGTCCACCGTGGAGTTTCGTCGATCGCG 843

Qy    922 AAGGGCAACCACGTTTTCATCGAAATGAACCCAGGTTGAGCACCGTGACT 981
          ||| |||||||||||||||||||||| |||||||||||||| ||||
Db    844 ACCGGCAACTTCTACTTCATGGAGAATGAATACCCGCACCCAAGTCGAGCATCCACA 903

Qy    982 GAAGAGTCACCGAGGTGGACCTGGTGAAGGCGCAGATGCGCTTGGCTTGCTGGTGCAACC 1041
          ||| | |||||||||||||||||||||||||| |||||||
Db    904 GAAATGATTACGGGACTGGACTTGATTGCGGAGCAGATTCGGATTGCCCAAGGCG----- 958

Qy   1042 TTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGC 1101
          |||||||||||||||||||||||||| |||
Db    959 ----AAGCGCTGCGCTTCCGGCAAGCCGATATTCAACTGCGGCGGCCAGATACCGGAATGC 1014

Qy   1102 CGCATCACCACGGAAGATCCAAACAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1161
          ||||||||||||| |||||||||| || ||||||||||||||||||||||||
Db   1015 CGTATCAATGCGGAAGATCCGGAATACAATTTCCGGCCGAATCCTGGCCGCCATTACAGGC 1074

Qy   1162 TACCGGCTCACCAGGCGGAGCTGGCGTTGGCCGTGTCTCTTGACGGTGCAGCTCCGTGGCGAA 1221
          ||| |||||||||||||||||||||||||||
Db   1075 TATTTACCGCCGGCGGGCGTTCGTGTCGTGATTCCCATGTTTATACCGACTACGAA 1134

Qy   1222 ATCACCGCACACTTTGACTCCATGCTGGTGAAAAATGACCTGCCGTGGTTCCGACTTTGAA 1281
          |||| ||||| |||||||||||| ||||||||||||||||||||||
Db   1135 ATTCCGGCCCTATTACGATTCGCTGATTGGCAAATTGATTGTCTCGGGGTGCAACACGGGAA 1194

Qy   1282 ACTGCTGTTGCTCGTGCACACGCGGCGGTTGGCTGAGTTCACCGTGTCTCGTGTTGCAACC 1341
          ||||||||||| |||||||||||||||||||||||||||||
Db   1195 GAGGCGATCGCGCGGATGCAGGGTGCTCGTGCGGGAATGCCGCCATCACCGGCTTGCCGACG 1254

Qy   1342 AACATTGGTTTCTTGCCGGTTCGGTGCTGCGTGGGGAAGAGAGACTTCACTTCCAAGGCGCATCGCC 1401
          ||| ||||||||||||||||| |||||||||||||||
Db   1255 ACCCTTAGTTTCCATCAGCTGATGTTGCAGATGCTTGCGAGATGTGCGAGTTCCTGCGGCGGGGGAACTCTAT 1314

Qy   1402 ACCGGATTTATCG 1414
          ||||||||
Db   1315 ACCAACTTTGTTG 1327
```

RESULT 12
US-08-074-121-1
; Sequence 1, Application US/08074121
; Patent No. 5767362
;
; GENERAL INFORMATION:
;   APPLICANT: Best, Elaine
;   APPLICANT: Knauf, Vic C.
;   TITLE OF INVENTION: Methods and Compositions for Modulating
;   TITLE OF INVENTION: Lipid Content of Plant Tissues
;   NUMBER OF SEQUENCES: 15
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road, Suite 100
;   CITY: Menlo Park
;   STATE: California
;   COUNTRY: US
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/074,121
;   FILING DATE: 08-JUN-1993
;   CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Rae-Venter, Barbara
;   REGISTRATION NUMBER: 32,750
;   REFERENCE/DOCKET NUMBER: 05938/043001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 854-5277
;   TELEFAX: (415) 854-0875
;   TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3077 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS

```
;    LOCATION:      800..1267
;    FEATURE:
;    NAME/KEY:      CDS
;    LOCATION:      1281..2626
US-08-074-121-1

Query Match       8.7%;    Score 302.2;    DB 1;    Length 3077;
Best Local Similarity  54.9%;    Pred. No. 6e-74;
Matches  714;    Conservative  0;    Mismatches  563;    Indels  24;    Gaps  5;

Qy       89  AAAAGATCTTGGTAGCAAACCGCGCGGAAATCGCCGTCCGTCCGTTCGGTGCAGCACTCG  148
             ||| |||||  ||| | |||  |||||||||||||  ||||||||  |||||||||
Db     1288  ATAAAATTGTTATTGCCAACCGGCCGGAGATTGCATTGCTATTCTTCGTGCCTGTAAAG  1347

Qy      149  AAACCGGTGCAGCCACGGTAGCTATTTACCCCGTGAAGACATCGGGGATCATTCCACCGCT  208
             || |  ||| |||  ||||||||||||||||||||||||||||||||||||||||||
Db     1348  AACTGGGCATCAAGACTGTCGCTGCGTGCACTCCAGGCGCGGATCGGCGATCGCCAGCGTAT  1407

Qy      209  CTTTTGCTTCTGAAGCTGTCCGCATTGGTACTGAAGGCTCACCAGTCAAGGCGTACCTGG  268
             ||||||||||||||||||||||||||||   ||||||||||||| || ||||
Db     1408  TACTGGCAGATGAAACGGGTCTGTATTGGGCCTGCCTGT---CAGTAAAAAGTTATCTGA  1464

Qy      269  ACATCGATGAAAATTATCGGTGCAGCTAAAAAGTTAAAGCAGATGGTATTTACCCGGGAT  328
             ||||  |||||||| ||||||||||||||||||||||||  |||||||||||||||
Db     1465  ACATCCCGGCAATCATCAGCGCCGCTGAAATCACCGGCGCAGTAGCAATCCATCCGGGTT  1524

Qy      329  ATGGCTTCCTGCTGAAAAATGCCCAGCTTGCCCGGCAGTGCGCGGAAAACGCATTACTT  388
             ||||  ||||||||||||||| ||||||||||||||||||||||||||||||||
Db     1525  ACGGGTTCCTCCTCCGGAGAACGCACTTTGCCCGGACGCTCCGGCTTTATCT  1584

Qy      389  TTATTGGCCCAACCCCAGAGAGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTAACCG  448
             ||||||||||| |  |||||  | |||||||||||||||||||||||||||||
Db     1585  TCATTGGCCCGAAAGCAGAAACCATTCGCCTGATGGCGACAAAGTATCCGCAATCGCGG  1644

Qy      449  CCGGGAAGAAGGCTGGTCTCTGCCAGTTTTGGCGGAATC------CACCCCGAGCAAAACA  502
             || ||||| ||||||||||||||||||||||||| ||||      ||||||||||||||
Db     1645  CGATGAAAAAAGGGGGCGTCCCTTGCGTACCGGGTTCTGACGGCCCGCGTGGGCGACGATA  1704

Qy      503  TCGATGACATCGTTAAAAGCGCTGGAGGCCAGACTTACCCCATCTTTGTAAAGGCAGTTG  562
             ||||||||  | ||||  ||||| ||||||||||||||||||||| ||||||||||||
Db     1705  TGGATAAAAACCGTGCCATTGCCATTGGTTATCCGGTGATTATCCAAAGCCTCCG  1764

Qy      563  CCGGTGGTGGCGGACGCGGGTATGGCGGTATGCCGGGTAGTGCCGGACGCTGAACTGGCCA  622
             |||||||||||||| |||||||||||||||| ||||||||||||||||||
Db     1765  GCGGCGGCGGCGGTGTGCGGTATGCGGTATGCGCGGGCGACGCTGAACTGGCACACCATCCA  1824

Qy      623  CAACAGAAGCATCTCGTGAAGCTGAAGCGGCATTCGGCAGCGGCATTCGGCTCGGTATATGTCGAAC  682
             ||| ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     1825  TCTCCATGACCCGTGCGGAGCGGAAAGCTGCTTTCAGCAACGATATGGTTTACATGGAGA  1884

Qy      683  GTGCTGTGATTAACCCCAGCACATTGAAGTGCAGATCCTTGGCGATCGACTGCGACTGAGAAG  742
             || ||||||||||||||||||||||||||| |||||||||||||||||
Db     1885  AATACCTGAAAAATCCTGCCACCGTTCAGTTCAGGTACTGGCTCGACGGTCAGGGCAACG  1944

Qy      743  TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAGTGTTCGAAA  802
             ||||||||||||||||| |||| |||||||||||||||||||||||||
Db     1945  CTATCTATCTGCGGAACGTGACTGACTGCTCCATGCAACGCGCCACGGGAAAGTGGTCGAAG  2004

Qy      803  TTGGCCAGCCAGCACAGCATTTGATGATCCGAGAACTGCCTGATCCGCATTTGTGCCGATGCAGTAA  862
             ||||||||||||||| |||  || ||||||||||||||||||||||||||||
Db     2005  AAGCGGCCAGCCAGCACCGGCCATTACCCCGGAACTGCGTCGCTCGCGGCGAACGTTGCGCCTA  2064

Qy      863  AGTTCTGCGCCTCCATTGGTTACCAGGGGCGCGGGAACCTGGAATTCTTGGTGATGAAA  922
             ||| || ||||||||||||||||||||||||||| |||||||||| ||
Db     2065  AAGCGGCCAGCACCGGCCATTACCCCGGAACTGCGGTGCAGGTGCTATCGCGTGCTGCCGTCGT  2121

Qy      923  AGGGCCAACCACGTTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTG  982
             ||||| || ||||||||||||||||| | ||| ||||||||||||
Db     2122  ACGGCGAGTTCTATTTCATCGAAATGAACAACCCGTATTCAGGTAGAACACCCGGTTACAG  2181

Qy      983  AAGAAGTCACCGAGCTGGACTTGAAGGCGGCAGATGCCGTGCTTGGTGTCGCAACCT  1042
             || |||||||| | |||||| ||||||||||| ||| ||||||||||||| |||||
Db     2182  AAATGATCACCGGGGGGGTTGACCTGACCGCATCAAAAGAACAGCTGCGTATCGCTGCCGGTCAACCGC  2241
```

```
Qy     1043  TGAAGGAATTGGGTCTGACCCAAGATAGATCAAGACCCACGGTGCGAGCACTGCAGTGCC  1102
             |||||  |||||||||||||  ||||||||||| |||| |||||||||  |||
Db     2242  TGTCG-------ATCAAGCAAGAAGAAGAAGTTCACGTTCGCGCGCCATGCGCGGTGGAATGTC  2292

Qy     1103  GGATCACCACGGAAGATCAAACAACACGGCTTCCGCCCAGATACGGAACTATCACCCGT  1162
             ||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db     2293  GTATCAACGCCGAAGATCCGAACA---CCTTCCTGCCAAGTCCGGGCAAAATCACCCGTT  2349

Qy     1163  ACCGCTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTGCAGCTCAGCTGCTGGCGAAA  1222
             |||||||||||||||||||||||   |||||||||||||||  ||||||||| ||||
Db     2350  TCCACCGCACCTGGCGTTTTGCGTACGTTGGGAGTCGTTCATATCTACGCGGGCTACACCG  2409

Qy     1223  TCACCGCACACTTGACTCATGCTCGTCGTGAAAATGACACCTGCCGTGGTTCCGACTTTGAAA  1282
             |||| |||||||||||||||||||||||| |||||||||||||||||||
Db     2410  TACCGGCGTACTATGACTCAATGATCGTTAAGCTGAGTAAGCTGATTTGCTACGGTGAAAACCGTGACG  2469

Qy     1283  CTGCTGTTGCTGTGCACAGGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCA  1342
             ||||||||||||||||||||||||||  || |||||||||||||||  ||||||
Db     2470  TGGGATTGCCCGACTGCGCCATGATGAAGAATGCGCTGCAGGAGCTGATCATCGACGGTATCAAAAACCA  2529

Qy     1343  ACATTGGTTTCTTGCGTGCGTTGCTGCGGGAAGAGAGACTTC  1383
             |||||  ||||||||||||||||||||  ||| ||||||
Db     2530  ACGTTGATCTCGAGATCCGCATCATGAATGACGAGAGACTTC  2570
```

```
RESULT 13
PCT-US94-06447-1
;  Sequence 1, Application PC/TUS9406447
; GENERAL INFORMATION:
;    APPLICANT:  Calgene, Inc.
;    TITLE OF INVENTION:  Methods and Compositions for Modulating
;    TITLE OF INVENTION:  Lipid Content of Plant Tissues
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weil, Gotshal & Manges
;      STREET:  2882 Sand Hill Road, Suite 280
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/06447
;      FILING DATE:  06-JUN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rae-Venter, Barbara
;      REGISTRATION NUMBER:  32,750
;      REFERENCE/DOCKET NUMBER:  CGNE-097/WO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 926-6200
;      TELEFAX:  (415) 854-3713
;      TELEX:
;  INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3077 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  800..1267
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1281..2626
PCT-US94-06447-1
```

Query Match   8.7%;   Score 302.2; DB 5; Length 3077;
Best Local Similarity   54.9%;   Pred. No. 6e-74;
Matches 714;   Conservative   0;   Mismatches 563;   Indels 24;   Gaps 5;

```
Qy    89 AAAGATCTTGGTAGCAAACCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCACTCG 148
          ||||| ||| | ||||| ||||||||||| ||||| ||||||| |||||||||||
Db  1288 ATAAAATTGTTATTGCCACCGGCGGAGATTGCATTGCGTATTCTTCGTGCCTGTAAAG 1347

Qy   149 AAACCGGTGCAGCCACGGTAGCTATTGCCGTGAAGATCGGGGATCATTCCACCGCT 208
          | |||||||| |||||| | |||||| ||||||||||||| |||||||||||||||
Db  1348 AACTGGGCATCAAGACTGTCGCTGCACTCCAGCGCGGATCCGATCTAAAACACGTAT 1407

Qy   209 CTTTTGCTTCTGAAGCTGCCATTGGTACTGAAGGGCTCACCAGTCAAGGGCTACCTGG 268
          ||||  ||||||||  |||| |||| |||||| ||||||||||||| ||||| |||
Db  1408 TACTGGCAGATGAAACGGTCTGTATTGCCCGTCCGT---CAGTAAAAAGTTATCTGA 1464

Qy   269 ACATCGATGAAATATCGGTGCAGTCAGAAAAAGTTAAAGCAGATGCTATTTACCCGGGAT 328
          ||||||||||||   || ||| |||||| ||||||||| |||| ||| ||||||||
Db  1465 ACATCCCGGCAATCAGCGCGCGCTGAAATCACCGGCCAGTAGCAATCCATCCGGGGTT 1524

Qy   329 ATGGCTTCCTGCTGTCGAAAATGCCCAGCTTGCCCGGAGTGCGCCGGAAAACGGCATTACTT 388
          ||||||| ||||| || | |||| |||||||| |||| ||||||||| ||||||||||
Db  1525 ACGGCTTCTCTCTCCCGAGAACGCCAACTTTGCCCGGAGCAGTTGAACGTCTCCGGCTTTATCT 1584

Qy   389 TTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCCGGTGATAAGTCTCGTGCGGTAACCG 448
          ||||||| |||| ||||||||| |||| |||||||||| ||||| |||||||||
Db  1585 TCATTGGCCCGAAAGCAGCAGAACCATTCGCGTGATGGGCGACAAAGTATCCGCAATCGCGGG 1644

Qy   449 CCGCGAAGAAGGCTGGTCTGCCAGTTTTGCGGGGAATC------CACCCCGAGCAAAAACA 502
          ||||||||||||| ||||||| ||||||| |||||||      |||||||||||||||||
Db  1645 CGATGAAAAAAGCGGGGCGTCCCTTCTGCGACCGGCCCGCCCGGGCGACGATA 1704

Qy   503 TCGATGACATCGTTAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTAAGGCAGTTG 562
          |||||||| ||||||||| ||||||||| |||||||||||||||||||||||||||
Db  1705 TGGATAAAAACCGTGCCATTGCCAATTGGTTATCCGTGATTATCAAAGCCTTCCG 1764

Qy   563 CCGGTGGTGGCGGACGCGGTATGCGCGTTTGTTCTTCACCTGATGAGCTTCCGCCAAATTGG 622
          ||||||||||||| ||||||||||| ||| || ||||||||||||| ||||||||||||
Db  1765 GCGCGGCGGCGGTCGCGGCGTATGCCGCGTAGTGCCGGCGCACGGCTGAACTGGCACAATCCA 1824

Qy   623 CAACAGAAGCATCTCGTGAAGCTGAAGCCGCATTCGGCGACGGTTCGGTATATGTCGAAC 682
          |||||||||||||| | |||| ||||| |||||| |||| |||||| ||||||||||
Db  1825 TCTCCATGACCGTGCCGGAAGCGGAAAGCTGCTTTCAGCAACGATATGGTTTACATGGAGA 1884

Qy   683 GTGCTGTGATTAACCCCAGCAGCACATTGAAGTCGAGATCCTTGGCGATCGCACTGGAGAAG 742
          ||||||| || |||||||||||||| | ||||| |||| ||||||||| ||||||||||
Db  1885 AATACCTGGAAAAATCCTCGCCACGTCAGGTACTGGCTGCGTCACCGAAAAGTGTCGAA 1944

Qy   743 TTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTGTCGAA 802
          ||||| |||| | |||||| || ||||| |||||||| ||||||||||||||||
Db  1945 CTATCTATCGGCGAACGTGACTGCTCCTCCATGCAACGCGCCACCAGAAAGTCGTGCGAAG 2004

Qy   803 TTGGCGCCAGCACAGCATTGGATCCAGAACTGCGTTGTGCGGATGCAGTAA 862
          ||||| ||| ||||||| |||||||| |||||||||||| |||||||||
Db  2005 AAGCGGCCAGCACCGGGCCATTACCCCGGACTGCCTCGCTACATCGGCGGAATGTC 2064

Qy   863 AGTTCTGCCGCTCCATTGGTTACCAGGGGCGGCGGAACCTGGAATTCTTGGTCGATGAAA 922
          |||||||| | | |||||||||||| | ||||||||||||| ||||||| ||||||||
Db  2065 AAGCGGCAGCACCGGGCATTACCCGGGCTATCCGGGTCATCCGGTCAGGTACTTTC---GAA 2121

Qy   923 AGGGCAACCACGTTTTCATCGAAATGAACCACGTATCCAGGTTGAGCACCGTGACTG 982
          |||| | || |||| ||||| ||||| |||| ||| |||||| |||| | ||| |
Db  2122 ACGGCGAGTTCTATTTCATCGAAATGAACCACGTATTCAGGTAGAGAACACCCCGGTTACAG 2181

Qy   983 AAGAAGTCACCGAGTGGACCTGGTGGAAGCGCAGATGCGCAGATGGCCGTGCTGGCTGTGCAACCT 1042
          |||||  |||||| |||| || | ||| ||||| |||||| ||||| ||||| ||||| ||
Db  2182 AAATGATCACCGCGGTTGACCTGACCGAAGAACAGCTGCGTATCGCTGCCGGTCAACCGC 2241

Qy  1043 TGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGTGCC 1102
          || |  || |||| | |||| ||| ||| |||||| | |||||||| ||| ||||||||
Db  2242 TGTCG-------ATCAAGCAAGAAGAGTTCACGTTCGCGGCCATGCGCGGTGAAATGTC 2292

Qy  1103 GCATCACCACGCGGAAGATCCAAACAACGGCTTCCGCGCCAGATACCGGAACTATCACCGCGCT 1162
```

```
Db  2293 GTATCAACGCCGGAGATCCGGACA---CCTTCCTGCCGTCTTGCTTTCCGGGCAAAAATCACCCGGTT 2349
          ||||| ||||||||||||| ||||   | |||||||||||| ||| ||||| || |||||||
Qy  1163 ACCGCTCACCAGCGGAGTCGCTGCGTTCGTGCTGCAGCTCGCGAGTCGAGGCGCAGCTCAGCCGAAA 1222

Db  2350 TCCACGCACCTGGCGGCGTTTGCGTTACGTTGCGGGCGACGTTCATATCTACGCGGCGGCTACACCG 2409
          | |||| || || |||| ||| ||| || ||| ||| |||||| || || ||||| || |||||
Qy  1223 TCACCGCACACTTTGACTCCATGCTCGTGTGAAAAATGACCTGCCGTGGTTCCGACTTTGAAA 1282

Db  2410 TACCGCCGTACTATGACTCAATGATCGATGCTGTGGCTGTCTCGTAACGGTGAAAAACCGTGACG 2469
          ||||| |||||||||| |||| |||||||| ||| ||||| ||||| ||||||||| || | |
Qy  1283 CTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCTCACCCGTCTCGGTGTTGCAACCA 1342

Db  2470 TGGCGATTGCCCGCACCATGAAGAATGGCGCTCGAGGAGCTGATCATCGACGCTGATCAAAACCA 2529
          ||||| |||| |||| |||| |||||| || |||| |||| ||||||||| ||||||||||||
Qy  1343 ACATTGGTTTCTTGGCGGCGTTGCTGCGGGAGGAGACTTC 1383

Db  2530 ACGTTGATCTGCGAGATCCGCATCATGAATGACGAGAACTTC 2570
```

```
RESULT 14
US-07-956-700B-1
;   Sequence 1, Application US/07956700B
;   Patent No. 5539092
;   GENERAL INFORMATION:
;     APPLICANT:  Robert Haselkorn and Piotr Gornicki
;     TITLE OF INVENTION:  Cyanobacterial and Plant Acetyl-CoA
;     TITLE OF INVENTION:  Carboxylase
;     NUMBER OF SEQUENCES:  116
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Arnold, White & Durkee
;       STREET:  321 No. 5539092th Clark Street
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60610
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy Disk
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII-DOS
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/956,700B
;       FILING DATE:  19921002
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Thomas E, No. 5539092thrup
;       REGISTRATION NUMBER:  33,268
;       REFERENCE/DOCKET NUMBER:  ARCD.058
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  1-312-744-0090
;       TELEFAX:  1-312-755-4489
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3065 base pairs
;       TYPE:  Nucleic acid
;       STRANDEDNESS:  Single
;       TOPOLOGY:  Linear
;     MOLECULE TYPE:  Oligonucleotide
US-07-956-700B-1
```

Query Match   8.5%;   Score 297; DB 1; Length 3065;
Best Local Similarity   54.4%;   Pred. No. 1.7e-72;
Matches 691;   Conservative   0;   Mismatches 565;   Indels 15;   Gaps 4;

```
Qy    85 TTCAAAAGATCTTGGTAGCAAACCGCGGCGAAATCGCGGTCCGTGCTTCCGTGCAGCA 144
          || ||| | | |||||||||| | ||||| |||| ||||||||| ||||||||||||
Db  1314 TTTGACAAAATATTAATTGCCAATCGGGGAGAATAGGCGCTGCGCTGCCAGTCTCTGCCTGT 1373

Qy   145 CTCGAAACCGGTCAGCCACCACGGTAGCTATTACCCCGTGAAGATCGGGGATCATTCCAC 204
          |||||||| || | ||||| | || ||| | |||||| ||||||||||| |||||||
Db  1374 GAGGAAAATGGGAATTGCGACGATTCGCAGTTCATTCGACTGTTGCACCGGAATGCTCTTCAT 1433
```

```
Db  2502 ACTGCTATTAACCGCATGATGAACGCGCCCTCAGGAATGCGCCATCACTGGATTACCACTACA 2561
         ||||| ||    |||          ||||      |||| |||     ||||   |||| ||||
Qy  1342 AACATTGGTTT 1352
         ||||||| ||
Db  2562 ACCATTGGGTT 2572
```

RESULT 15
US-08-476-537-1
; Sequence 1, Application US/08476537
; Patent No. 5756290

GENERAL INFORMATION:
  APPLICANT: Robert Haselkorn and Piotr Gornicki
  TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
  TITLE OF INVENTION: Carboxylase
  NUMBER OF SEQUENCES: 116
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Arnold, White & Durkee
    STREET: 321 No. 5756290th Clark Street
    CITY: Chicago
    STATE: Illinois
    COUNTRY: USA
    ZIP: 60610
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy Disk
    COMPUTER: IBM PC Compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: ASCII-DOS
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/476,537
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/956,700
    FILING DATE: 10/21/92
  ATTORNEY/AGENT INFORMATION:
    NAME: Thomas E. No. 5756290thrup
    REGISTRATION NUMBER: 33,268
    REFERENCE/DOCKET NUMBER: ARCD:058
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 1-312-744-0090
    TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 3065 base pairs
    TYPE: Nucleic acid
    STRANDEDNESS: Single
    TOPOLOGY: Linear
  MOLECULE TYPE: Oligonucleotide
US-08-476-537-1

Query Match   8.5%; Score 297; DB 1; Length 3065;
Best Local Similarity 54.4%; Pred. No. 1.7e-72;
Matches 691; Conservative 0; Mismatches 565; Indels 15; Gaps 4;

```
Qy    85 TTCAAAAGATCTTGGTAGACAAACGCGGCGGCGAAATCGCGGGTCCGGTCGTTCCGTGCTGCAGCA 144
         || |||      |    | |  ||||      |  |||||  || ||| ||||  |  | |||||
Db  1314 TTTGACAAAATATTAATTGCCAATGCGGGAGAAATAGCGCGTCTGCGCATTCTCCGCGCCTGT 1373

Qy   145 CTCGAAACCGGTGCGAGCCACGGTAGCTATTACCCCCGTGAAGATCGGGGGATCATTCCAC 204
         |  ||| |   |||    |  |    || |||| || |     |    || |  || |||||
Db  1374 GAGGAAATGGGGATTGCGACGATCGCAGTTCATTGGTGACCGGAATGCTCTTCAT 1433

Qy   205 CGCTCTTTTGCTTCTCCGAAGTCGTCCGCATTGGTACTGAAGGCTCACCAGTCAAGGCGGTAC 264
         ||| |   ||  || | || |  |  ||||         |||        |  | ||   |||
Db  1434 GTCCAACTTGCTGACGAAGCGGTTTGTATTGGC---GAACCTGCTAGCGCTAAAAGTTAT 1490

Qy   265 CTGGACATCGATGAATTATCGGTCGAGCTGAGCTAAAGCAGATGCTATTACCCG 324
         |   || |||||| |  |      | | |        | |  || |||| |
Db  1491 TTGAATATTCCCAATATTATTGCTCGCGGCTTTAACGCGCAATGCCAGTTGCTATTCATCCT 1550

Qy   325 GGATATGGCTTCCTGTCTGTGAAAATGCCCGAGCTGCGCGGAAAACCGGCATT 384
         |||||||||||   ||||||||||||||||    ||||| |||||| |  |||
Db  1551 GGGTATGGCTTTTTATCTGAAAATGCCAAATTGCGGAAATCTGTCGTGACCATCACATT 1610

Qy   385 ACTTTTATTGGCCCAACCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGTGCGGTA 444
         ||| |||| |||||| ||| |||||   |||||  |||     ||   || |||| |||
Db  1611 GCATTCATTGCGCCCCACCCCAGAGTCGCCTCAGGGGGACAAATCCACTGCCAAG 1670

Qy   445 ACCGCGCGGAAGAAGGCTGGTCTCGACCAGTTTGGCGG--AATCCACCCCGAGCAAAAACA 502
         ||| | || || ||||||  |  ||| |||||| |    ||| ||    | |||  ||||
Db  1671 GAAACCATGCAAAAAGCTGTGGTGTACCGACAGTACCCGGGTAGTGAAGGTTTGGTAGAGACA 1730

Qy   503 TCGATGACATCGTTAAAAGCCTGAAGCCA-GACTTACCCCATCTTGTAAAGGCAGTT 561
         || ||||| || |  |||||||||  ||| ||  ||||   |  || |  || |  |
Db  1731 GAGCAAGAAGGATTAGAACTGGCGAAAGATATTGGCTACCCAGTGATGATGATCAAAGCCACG 1790

Qy   562 GCCGGTGGTGGCGGACGGCGTATGCGCTTTGTGTTTCTTCACCTGATGAGCTCCGCAAATTG 621
         || ||||||||||| || |||||||  | |||   | || |  || |||||   |||||||
Db  1791 GCTGGTGGTGGCGGCGCGGGGTATGCGCGACTGTGCGCATCGCCAGATGAATTTGTCAAACTG 1850

Qy   622 GCAACAAGAGCGATCTCGTGAAGCTGGCGCATTCGGACGCGGTTCGGTTATATGTCGAA 681
         ||| ||| ||| |||  || ||||| |||  |  |   | |||| ||| ||||| ||||
Db  1851 TTCTTAGCCGCCAAGGTGAAGCTGTGTGCAGCCTTTGGTAATGTGGCGTTTTATATAGAA 1910

Qy   682 CGTGCTGTGATTAACCCCCAGCACAATGAAGTGCAGATCCTTGGCGATCCGACTGAGAA 741
         | | || || ||   |   || | || |  || |||  |  |||| ||    ||| ||
Db  1911 AAATTTATTGAACGTCCGCCCACATTGAATTTCAAATTTTGGCTGATAATTACCGCAAT 1970

Qy   742 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 801
         |    ||||| |     |  |||  |||| | | ||||||| ||||||||||| | |||
Db  1971 GTGATTCACTTGGGTGAGAGGGATTGCTCAATTCAGCGTCGTAACCAAAAGTTACTAGAA 2030

Qy   802 ATTGCGCCAGCACAGCAGCATTTGGATCCAGAACTGCGCTGATCGATCGCATTGTGCGGATCAGTA 861
         | ||||||||||||||  | |||||| ||||  || ||  ||| ||
Db  2031 GAAGCCCCCAGCCAGCCTTGGACTCAGACTGCGCAGGGAAAAATGGGACAAGCGGCGGTG 2090

Qy   862 AAGTTCTGCGCCTCCATTGGTACCAGGCGGCGGGAACCGTGGAATTCTTGGTCGATGAA 921
         |||  | |||| | |  ||| |||||||||||| ||||| | |  || ||  ||||  |
Db  2091 AAAGGCGGCTCAGTTTATCAATTACCGGGCGGCAGTGCTACTATCGAGTTTTGCTAGATAGA 2150

Qy   922 AAGGGCAACCACCAGTTTCTACTTTATGGAGATGAACAACATCCCGTGACT 981
         || ||| | || |  || |   || |||||||||||||  || |  |||  |||
Db  2151 TCCGGTCAGTTTTACTTTATGGAGATGAACAAGTTCAAGTAGAACATCCCAACT 2210

Qy   982 GAAGAAGTCACCGAGGTGGACCGAGGTGAAGCGCAGAAGCGGCAGATGCGCTTGCCTGCAACC 1041
         ||| |  | || ||  ||| |   |||||| |||| ||| | ||||| |||  |||  |||||--
Db  2211 GAGATGGTTACTGGACGTGGAGTTTATTGGTTGAGCAAATCAGAAATTGCCCAAGGGGAAA-- 2268

Qy  1042 TTGAAGGAATTGGGTCGACTGAAGATAAGATCAAGACCCCACGGTGCAGCACTGCAGTGC 1101
         |||||      |    ||||  |   |  ||||||| || |  ||||  | | | |||| |
Db  2269 ------GACTTAGACTAACTCAAGACCAAGTTTTTACGGTGTTCATGCGGATCGAATGT 2321

Qy  1102 CTGCATCACCAGCGAAGATCCAAGAAACAACGGCTTCCGCCCAGATACCGAACTATCACCCGCG 1161
         | ||| || ||| |||||| ||||  | | || ||||||||||  |||    |||| || ||||
Db  2322 CGCATCAATGCCGAAGACCCAGACCACGATTTCCGCCCAGCACACCCCAGCGCCATTAGCGGT 2381

Qy  1162 TACCCGTCACCAGGCGGAGCTGGCGTTCGTCTTGACGGTCAGCTCAAGTCGGTGGCGAA 1221
         || |  | |  ||||   |||| ||| |||  ||||| || ||   ||| | ||||||
Db  2382 TATCTTCCCCCTGGCGCCCTCGGCGTGCGGGATTGACTCCCACGTTTACACGGATTACCAA 2441

Qy  1222 ATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCGCGGTTCCGACTTTGAA 1281
         || || ||||| |||||| |  | |  |||||| ||  | |||  || |  ||| |||
Db  2442 ATTCCGCCCACTACGACGATTCCTTAATTGGTAAATGTAAATTGATCGTTGCGGCGCGCT 2501

Qy  1282 ACTGCGTTGCTGCGTGCACAGCGCGCGCGGTTCGCAGCGGCGGCGGAATGCGGCATT 1341
```

Wed Apr 9 09:38:21 2003

us-09-974-973-1.rni

Job time : 152 secs

```
Db 1551 ||| |||||||| | |||||||||||||| |||| ||| |||||  ||||| 1610
         GGGTATGGCTTTTATCTGAAAATGCCAAATTGCGGAAATCTGTGCTGACCATCACATT
Qy  385 ACTTTTATTGGCCCAACCCCAGAGTTCTTGATCTCACCGGTGATAAGCTCTCGTGCGTA 444
         |||| |||||||||||||| |||| || ||||| |||||| |||||||
Db 1611 GCATTCATTGGCCCCACCCCAGAGAGCTATCCGCCTCATCGGGGACAAATCCACTGCCAAG 1670
Qy  445 ACCGCCGCAAGAAGAGGCTGTCTGCCAGTTTGGCGG--AATCCACCCGAGCAAAAACA 502
         |||||| ||||||||| ||||||||| ||| |||||| |||  |||
Db 1671 GAAACCATGCAAAAAGCTGTGTACCGACAGTACCGGGTAGTGAAGGTTTGGTAGAGACA 1730
Qy  503 TCGATGACATCGTTAAAAGCCTGAAGGCCA-GACTTACCCCATCTTTGTAAAGGCAGTT 561
         |||| |||||||||| ||||| |||| |||| ||||
Db 1731 GAGCCAAGAAGGATTAGAACTGGCGAAAGATATTGGCTACCCAGTGATGATCAAGCCACCG 1790
Qy  562 GCCGGTGGTGGCGACGCGTATGCGCTTTGTTTCTTCACCTGATGAGCTCCGCAAATTG 621
         ||| |||||||| |||||| ||| |||||| |||||||||| ||| |||
Db 1791 GCTGGTGGTTGGCGCCGGGGTATGCGACTGGTGCGATCGCCAGATGAATTTGTCAAACTG 1850
Qy  622 GCAACAGAGCATCTCGTGAAGCGGCATTCGGCCGACCGGTTCGGTATATGTCGAA 681
         ||||||||| |||||||||||| ||| || || |||| |||||
Db 1851 TTCTTAGCCGCCCAAGGTGAAGCGTGTGCAGCCCTTTGGTAATGCTGGCCTTTTATATAGAA 1910
Qy  682 CGTGCTGTGATTAACCCCCGACACATTGAAGTGCAGATCCTTGGCGATCGCACTGGAGAA 741
         ||||||||||  || ||| ||||||||
Db 1911 AAATTTATTGAACGTCCGCGCCACCATTGAATTTCAAATTTGGCTGATAATTACGGCAAT 1970
Qy  742 GTTGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAA 801
         ||||| |||| ||| || || ||| ||||||||||
Db 1971 GTGATTCACTTGGGTGAGAGGGATTGCTCAATTCAGCGTCGTAACCAAAAGTTACTAGAA 2030
Qy  802 ATTGCGGCCAGCACAGCATTTGGATCCAGAACTGCCGTGATCGCATTTGTGCGGATGCAGTA 861
         |||| || || |||||||||| |||| || |||
Db 2031 GAAGCCCCCAGCCCAGCCTTGGACTCAGGACCTAAGGGAAAAAATGGGACAAGCGCGGGTG 2090
Qy  862 AAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA 921
         || ||||||| |||||||| ||| ||| ||||
Db 2091 AAAGCGGGCTCAGTTTATCAATTACGCCCGGGCAGTACTATCGAGTTTTTGCTAGATAGA 2150
Qy  922 AAGGGCAACCACGTTTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACT 981
         ||| |||||| |||||| ||||||||| |||| |||
Db 2151 TCCGGTCAGTTTTACTTTATGGAGATGAACACCCCGATTCAAGTAGAACATCCCGTAACT 2210
Qy  982 GAAGAAGTCACCGAGGTGGACCTGGAAGGCGCAGATGCGCTTGGCTGCTGCTGCAACC 1041
         ||| ||| ||| ||||| |||||| ||| |||
Db 2211 GAGATGGTTACTCGGAGTGGATTTATTGGTTGAGCAAATCAGAATTGCCCAAGGGGAAA-- 2268
Qy 1042 TTGAAGGAATTGGGTCTGACCAAGATAAGACTCAAGATAAGACGGCTTCCGCCCAGATACCGGAACTATCACCGCG 1101
         |||| ||| |||||||||| |||| |||
Db 2269 -------GACTTAGACTAACTCAAGACCCAAGTAGTTTTACGCGGTCATGCGATCGAATGT 2321
Qy 1102 CGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGGAACTATCACCGCG 1161
         |||||| |||||||||| ||||| |||| |||
Db 2322 CGCATCAATGCCGAGAGACCCGAGAGACCACCGATTTCCGCCCCAGCACCCGGACGCATTAGCCGGT 2381
Qy 1162 TACCGCTCACCAGGCGAGCTGGCCTGGCCTTCGTTCGTCGGTGCAGCTCAGCTCGGTGGCGAA 1221
         ||||||| |||||||| ||||||  ||| |||| |||
Db 2382 TATCTTCCCCCCGCGGCGGGCGTGCCGAATTGACTCCCACGTTTACACGGATTACCCA 2441
Qy 1222 ATCACCGCACACTTTGACTCCATGCTGGTGAAAAATGACCTGCCCGTGGTTCCGACTTTGAA 1281
         ||| ||||| |||||| ||||| |||
Db 2442 ATTCCGCCCTACTACGATTCCTTAATTGGTAAATTGATCGTTTGGGGCCCTGATCGCGCT 2501
Qy 1282 ACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACC 1341
         |||||||| ||||||| |||| || ||||
Db 2502 ACTGCTATTAACCGACCATGAAAACGCGCCCTCAGGGAATGCCCATCACTGGATTACCTACA 2561
Qy 1342 AACATTGGTTT 1352
         ||||||||||
Db 2562 ACCATTGGGTT 2572
```

Search completed: April 5, 2003, 10:03:09